DTIC FILE COPY

AD-A217 013

BRF ANALYSIS AND IMPROVEMENT

JAN 23 1990

OPERATIONS ANALYSIS DEPARTMENT

NAVY FLEET MATERIAL SUPPORT OFFICE

Mechanicsburg, Pennsylvania 17055

Report 169

BRF ANALYSIS AND IMPROVEMENT

Project Number
9321-E65-8207

Report 169

Submitted by:
R. E. MAITLAND
Operations Research Analyst

Approved by:
K. T. ADAMS, LCDR, SC, USN
DIRECTOR, OPERATIONS ANALYSIS DEPARTMENT

H. M. HARMS, CAPT, SC, USN
COMMANDING OFFICER
NAVY FLEET MATERIAL SUPPORT OFFICE

DATE DEC 26 1983

ABSTRACT

The Best Replacement Factor (BRF) is used primarily in Coordinated Shipboard Allowance List (COSAL) and load list requirements determination. Recently, many inventory problems including stock shortages, stock excesses, and churn have been blamed on the BRF. To identify what problems are associated with the current method used to compute the BRF, we took a multi-faceted approach. First, we discussed the two forecasting methods available under Resystemization and then analyzed the volatility and trend of the current BRF. Casualty Report (CASREP) items and problem equipments were included in the analysis. We then examined alternative methods to forecast the BRF and compared them with the current methods. The best alternative was selected based on its performance in terms of stability, accuracy with respect to bias and extremes, and the magnitude of errors. Performance was evaluated separately for items identified as readiness contributors. Test COSALs were built with the current BRFs and the proposed BRFs to evaluate the impact of the change.

We recommend using the exponential probability distribution to adjust the Technical Replacement Factor (TRF) during the development period, and after the development period, exponential smoothing with a data windsorizing system to handle extreme values.

TABLE OF CONTENTS

EXECUTIVE SUMMARY

1. Background. The Best Replacement Factor (BRF) is used primarily in Coordinated Shipboard Allowance List (COSAL) and load list requirements determination. Recently, many inventory problems including stock shortages, stock excesses, and churn have been blamed on the BRF.

2. Objective. To identify and investigate problems associated with the current method used to compute the BRF. After these problems have been identified and investigated, develop alternatives to improve the BRF forecast.

3. Approach. The first phase began with an analytical discussion of Exponential Smoothing and the Ratio Method (the two methods available under Resystemization). Next, we measured current BRF volatility over time by examining BRF changes and trends. Casualty Report (CASREP) requisitions were then analyzed. From the analysis of CASREP items, we identified and examined items that appeared to have major support problems. Finally we examined a single weapon system (PHALANX) in an attempt to identify specific BRF problems.

Taking the information gained in the first phase, we developed forecast methods to correct/improve the BRF. These new methods, as well as the current methods, were ranked based on their performance with respect to all items and readiness contributors. Performance was related to stability, accuracy with respect to bias and extremes, and magnitude of the errors. To evaluate the impact of any change, we compared the best method to the Benchmark (current forecast method) COSAL computed with Modified Fleet Logistics Support Improvement Program (MOD-FLSIP) rules.

4. Findings. BRF instability and trend are primarily related to the subjective portion of the forecast (Technical Replacement Forecasts (TRFs), ordnance freeze, manual changes and overrides). Eliminating the TRF, ordnance freeze,

manual changes and overrides from the forecast produced more stable forecasts that trended closer to the actual Navy Maintenance and Material Management (3M) average usage rates computed from all available data.

Using the exponential probability distribution to adjust the TRF during the development period, and exponential smoothing, with a data windsorizing system to handle extreme values after the development period, provided the best forecast with respect to stability, bias and extremes, and the magnitude of the errors. These results were consistent for all items and the subset of items classified as the readiness contributors.

The proposed method appears to slightly reduce COSAL range and effectiveness. However, we believe the improved stability and accuracy will provide long term benefits in reduced COSAL churn which tends to create long supply and/or outfitting deficiencies.

5. Recommendation. We recommend using the exponential probability distribution to adjust the TRF during the development period. After the development period, use exponential smoothing with a data windsorizing system to handle extreme values.

I. INTRODUCTION

The Best Replacement Factor (BRF) is used primarily in Coordinated Shipboard Allowance List (COSAL) and load list requirements determination. Recently, many inventory problems, including stock shortages, stock excesses, and churn have been blamed on the BRF. Some concerns are unrelated to the method used to compute the BRF. Poor reporting by activities, bad configuration data, and manual changes can lead to distorted forecasts. Other concerns are related directly to the current method used to compute the BRF. It has been suggested that using different data sources when Navy Maintenance and Material Management (3M) data is unavailable and using the current forecast method (Exponential Smoothing with a weight of .4) which reacts quickly to changes, create an unstable forecast.

This concern has grown since Readiness Based Sparing (RBS) models were approved for use in conjunction with the COSAL Modified Fleet Logistics Support Improvement Program (MOD-FLSIP) model. Under MOD-FLSIP rules, an item's requirements are based only on its expected demand (BRF times the item's activity population). Changes in an item's BRF only affect that item's requirements. With an RBS model, requirements are determined based on the relative demands and prices of all items being optimized; i.e., items compete against each other based on their expected demand and cost. A change in one item's BRF can affect not only its own requirements but those of other items as well As a result, the Ratio Method was proposed as a more stable alternative for Resystemization. The Ratio Method is the ratio of the cumulative 3M usage to the cumulative 3M population. This method is not yet approved for use, nor has it been shown to provide more cost-effective results.

Therefore, our initial phase of this study was to determine if any of these concerns are valid. After identifying and analyzing problems associated with the present BRF, we then propose and evaluate methods to correct/improve the forecast.

II. APPROACH

In this section, we describe the method, data and performance measurements used in our analysis

A. METHOD.

We begin the first phase with an analytical comparison of Exponential Smoothing and the Ratio Method, the two methods available under Resystemization. Next, we measure BRF volatility over time by examining BRF changes and trends. CASREP items are analyzed to identify and examine problem items. Finally, we analyze items on a sample problem equipment (PHALANX) in an attempt to identify specific BRF problems. Throughout the data analysis, we include an analytical discussion comparing what we observed with the theory.

Using the information gained in the first phase, forecast methods are developed to improve the BRF. These methods are ranked based on their performance with respect to all items and only those items identified as readiness contributors. Performance for items on two selected equipments is also analyzed. To measure the impact of any changes, the best method is compared to the benchmark COSAL (computed with the current forecast method and with MOD-FLSIP rules) for several test ships.

B. DATA.

The Naval Sea Logistics Center (NSLC) provided us with their 10 year BRF history file (1978-1987) and Casualty Report (CASREP) requisitions for the period 1 January 1986 to 31 December 1987. The 10 year history file contains the BRF and item demand and average population for 3M, Mobile Logistics Support Force (MLSF), and system data. We obtained PHALANX data from CACI.

These data were used to analyze potential BRF problems. The BRF history file was also used to compute BRFs using alternative procedures.

To measure the impact of any BRF change, Navy Ships Parts Control Center (SPCC) provided us with candidate files from the July 1988 COSAL extract for the following ships:

- CG 24 - USS REEVES
- DDG 6 - USS BARNEY
- DDG 16 - USS STRAUSS
- DDG 993 - USS SPRUANCE
- FFG 13 - USS SPRAGUE
- LPD 14 - USS TRENTON

NSLC provided us with 3M demand data from 1 August 1988 to 10 November 1988 to evaluate the test COSALs. NSLC also provided Weapon System File (WSF) extracts for two equipments; the PHALANX (Allowance Parts List (APL) number 006090052) and an engine (APL number L665360264). These data were used to more closely evaluate BRFs on sample equipments. To determine the impact on readiness, NSLC provided a list of Federal Supply Classifications (FSCs) to define which items they considered potential readiness contributors.

C. PERFORMANCE MEASURES.

Model range and requisition effectiveness, in conjunction with cost, was used to measure the impact on the COSAL. Model range effectiveness is the percent of COSAL candidate items with demand that was stocked by MOD-FLSIP. Model requisition effectiveness is the percent of requisitions for COSAL candidate items that was satisfied using the COSAL quantity.

BRF forecast performance was ranked based on stability, accuracy with respect to bias and extremes, and accuracy with respect to the magnitude of the forecast errors. The criteria used to rate the forecast methods is varied and at times contradictory. That is, the forecast method selected may or may not be the most stable or accurate, but it will provide the best combination of stability and accuracy given the alternative methods. It is important to note that all the measures look in hindsight to see how the yearly BRFs compare to the average 3M rate over time. The average BRF is based on 5-10 years data depending on the history available for the item.

The following paragraphs explain the forecast performance measures in detail.

1. Stability. We measured stability in terms of the number of times (years) the forecast is within 95% confidence limits of the average item 3M usage rate (computed from the 10 year history file). The more stable a forecasting method is, the more often its forecast will be within these limits.

$$UL = \bar{X} + t\left(\frac{S}{\sqrt{N}}\right)$$

$$LL = \bar{X} - t\left(\frac{S}{\sqrt{N}}\right)$$

where

UL = the upper confidence limit

LL = the lower confidence limit

$\bar{X}$ = the average usage rate for the n^{th} item computed as $\frac{1}{N}\sum_{i=1}^{N}$ 3M usage rate_i

N = the number of observations for the n^{th} item

S = the standard deviation for the n^{th} item computed as

$$\left[\frac{1}{N-1}\sum_{i=1}^{N}(\text{3M usage rate}_i - \bar{X})^2\right]^{\frac{1}{2}}$$

t = t-statistic used to define the 95% confidence limits

2. <u>Accuracy with Respect to Bias and Extremes</u>. We measured accuracy with respect to bias and extremes by the Mean Error (ME). The ME measures how forecasts err. A negative ME indicates a forecast tends to overforecast, while a positive ME denotes a tendency to underforecast. When an overforecast occurs, the item has a better opportunity to be stocked (or stocked with a larger quantity) and there is less risk of stockout (the part is likely to be available); however, investment may be larger than required (excess material may be bought). With an underforecast, the opposite is true (i.e., the item will have a worse opportunity to be stocked, there is a higher risk of stock-out, but less excess material will be bought). Forecast performance will be ranked by the size of the overforecast. Small overforecasts are ranked highest. Large underforecasts are ranked lowest.

$$\text{ME} = \frac{1}{N} \sum_{i=1}^{N} (\text{3M usage rate}_i - F_i)$$

where

ME = the mean error for the n^{th} item

N = the number of observations for the n^{th} item

F_i = the i^{th} forecast for the n^{th} item

3M usage rate_i = the actual 3M experienced rate for the n^{th} item during the i^{th} observation

3. <u>Accuracy with Respect to Magnitude of the Errors</u>. We measure accuracy with respect to the magnitude of the forecast errors by the Mean Square Error (MSE). Here we are interested in how close the forecast is to the 3M usage rate. Consequently, forecast performance is ranked on the ability to forecast precisely (with small errors).

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (3M \text{ usage rate}_i - F_i)^2$$

where

MSE = the mean square error for the n^{th} item

N = the number of observations for the n^{th} item

F_i = the i^{th} forecast for the n^{th} item

3M usage $rate_i$ = the actual 3M experienced rate for the n^{th} item during the i^{th} observation

III. FINDINGS

This section is comprised of two parts. In the first segment, we will identify and analyze problems associated with the current BRF. In the next section, we will propose alternatives for improving the BRF.

A. BRF PROBLEM ANALYSIS.

To determine what problems exist, and if the concerns mentioned previously are valid, we took a multi-faceted approach. First, we examined the two forecasting methods available under Resystemization; i.e., Exponential Smoothing and the Ratio Method. Next, we analyzed the volatility and trend of the current BRF. Then, we examined CASREP requisition data. Using the CASREP data, we identified and investigated problem items. Finally, we analyzed items on a problem equipment (the PHALANX data set).

1. Exponential Smoothing Versus the Ratio Method. The current BRF and the Ratio Method use different techniques to transition from the TRF to an experience-based demand rate. They can also use different data sources. For complete details see APPENDIX B. Our discussion focuses on the techniques used to phase observations into the forecast with an emphasis on the degree of adaptability and stability of the methods. The tradeoff between adaptability

and stability is usually controlled by the amount of influence or weight placed on recent observations versus older observations. APPENDIX C provides a detailed explanation of how these weights are developed for each method.

Exponential Smoothing is a moving average technique. The name is derived from how the influence or weight placed on an observation changes over time; i.e., it declines exponentially. For example, viewing TABLE I we see that the weight of the first observation (viewed for the first time in year two) declines from .4 to .007 over a nine year period using a smoothing weight of .4, and declined from .1 to .043 using a smoothing weight of .1.

The degree of adaptability and stability for this forecast method is determined by the smoothing weight (α). With a higher weight, more emphasis is placed on the more recent observation than on the older observations; therefore, the forecast adapts quicker to changes than with a lower weight, but is less stable. A lower weight places more emphasis on older observations than the more recent observations. Consequently, the forecast does not react as quickly as with a higher weight (less adaptable), but is more stable. This can be seen by viewing TABLE I. TABLE I provides a comparison of the emphasis placed on observations for up to a 10 year period, for high (.4) and low (.1) smoothing weights. For example, in the 10^{th} year, 64% (.24 + .4) of the high smoothing weight forecast is from the two most recent observations compared to 19% (.09 + .1) for the low smoothing weight forecast. The oldest two values (TRF and first observation) account for only 1.7% (.01 + .007) of the high smoothing weight forecast compared to 41% (.367 + .043) of the low smoothing weight forecast. Thus, the lower smoothing weight takes longer to "washout" the effects on the initial forecast (TRF).

TABLE I

EXPONENTIAL SMOOTHING WEIGHTS
WEIGHT = .4

YEAR	TRF	OBS 1	OBS 2	OBS 3	OBS 4	OBS 5	OBS 6	OBS 7	OBS 8	OBS 9
Y1	1.000									
Y2	0.600	0.400								
Y3	0.360	0.240	0.400							
Y4	0.216	0.144	0.240	0.400						
Y5	0.130	0.086	0.144	0.240	0.400					
Y6	0.078	0.052	0.086	0.144	0.240	0.400				
Y7	0.047	0.031	0.052	0.086	0.144	0.240	0.400			
Y8	0.028	0.019	0.031	0.052	0.086	0.144	0.240	0.400		
Y9	0.017	0.011	0.019	0.031	0.052	0.086	0.144	0.240	0.400	
Y10	0.010	0.007	0.011	0.019	0.031	0.052	0.086	0.144	0.240	0.400

WEIGHT = .1

YEAR	TRF	OBS 1	OBS 2	OBS 3	OBS 4	OBS 5	OBS 6	OBS 7	OBS 8	OBS 9
Y1	1.000									
Y2	0.900	0.100								
Y3	0.810	0.090	0.100							
Y4	0.729	0.081	0.090	0.100						
Y5	0.656	0.073	0.081	0.090	0.100					
Y6	0.590	0.066	0.073	0.081	0.090	0.100				
Y7	0.551	0.059	0.066	0.073	0.081	0.090	0.100			
Y8	0.478	0.053	0.059	0.066	0.073	0.081	0.090	0.100		
Y9	0.430	0.048	0.053	0.059	0.066	0.073	0.081	0.090	0.100	
Y10	0.367	0.043	0.048	0.053	0.059	0.066	0.073	0.081	0.090	0.100

The relationship between the size of the smoothing weight and stability becomes murky if distortion occurs; i.e., we observe unusual values. A distorted observation will have a greater initial impact on the forecast with a larger weight than with a lower weight, but the distortion will also be filtered out more rapidly with the higher weight than the lower weight. For example (viewing TABLE I), an observation viewed for the first time in year two accounts for 40% of the higher weight forecast as compared to 10% for the lower weight forecast. However, after four years (year six), that same observation will now account for more of the low weight forecast than the high weight forecast (6.6% versus 5.2%). By the 10^{th} year, that observation will have almost no influence on the high weight forecast (.7%) as compared to 4.3% for the low weight forecast.

Exponential Smoothing is a widely used technique for many reasons. First, it has been shown to provide viable results as compared to other methods in numerous studies. Second, it is relatively easy to understand and use. Finally, it only requires two data points to compute a forecast; i.e., the old forecast and the most recent observation. Consequently, Exponential Smoothing is the method most widely used for large inventory systems.

The main problems associated with Exponential Smoothing are developing a starting point and determining the smoothing weight. Smoothing weights are usually chosen analytically based on the desired tradeoff between stability and adaptability. The starting point is usually the first observation or an average of initial observations over time. As can be seen by TABLE I, the current policy (Exponential Smoothing with a weight of .4 and the TRF as a starting point) places a large influence on the TRF. The TRF always has more influence on the BRF than the first usage rate observed and almost as much weight as the second observed usage rate. It takes five years of data to reduce the TRF's impact on the forecast to less than 10%. Lowering the smoothing weight places an even larger influence on the TRF. After 10 years, the TRF still accounts for 37% of the BRF. More weight is given the TRF than the four most recent observations combined (six through nine). Since TRFs are only a best guess, this may cause a problem with the current BRF.

The Ratio Method was developed by the NSLC as an alternative to the current method. With this method, the observations are not smoothed into the

TRF. The TRF is the forecast (the BRF) until either two demands have or should have occurred. Then, the TRF is discarded, and the Ratio Method forecast becomes the BRF. The Ratio Method divides total accumulated demand by total population. This method claimed to weigh each observation equally and thus improve BRF stability. However, as shown in APPENDIX C, this is not necessarily the case. The weight of a yearly observation is dependent on changes in the population size used to compute the usage rate. If population is constant, then all observations are weighted equally. If population varies, the weight of each observation is determined by the size of the population generating that observation. For analytical purposes, we doubled the population size each year and halved the population size each year, for up to a 10 year period (see TABLE II). When population size increases, the forecast becomes more adaptable and less stable; and, when population size decreases, the forecast becomes less responsive and more stable. Population sizes used to calculate usage rates change yearly. They increase and decrease from year to year. Therefore, the degree of adaptability and stability changes from year to year, is beyond the control of the user, and is unknown.

TABLE II

RATIO METHOD WEIGHTS ASSUMING CONSTANT POPULATIONS

YEAR	OBS 1	OBS 2	OBS 3	OBS 4	OBS 5	OBS 6	OBS 7	OBS 8	OBS 9	OBS 10
Y1	1.000									
Y2	0.500	0.500								
Y3	0.333	0.333	0.333							
Y4	0.250	0.250	0.250	0.250						
Y5	0.200	0.200	0.200	0.200	0.200					
Y6	0.167	0.167	0.167	0.167	0.167	0.167				
Y7	0.143	0.143	0.143	0.143	0.143	0.143	0.143			
Y8	0.125	0.125	0.125	0.125	0.125	0.125	0.125	0.125		
Y9	0.111	0.111	0.111	0.111	0.111	0.111	0.111	0.111	0.111	
Y10	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100

RATIO METHOD WEIGHTS ASSUMING POPULATIONS DOUBLE EACH YEAR

YEAR	OBS 1	OBS 2	OBS 3	OBS 4	OBS 5	OBS 6	OBS 7	OBS 8	OBS 9	OBS 10
Y1	1.000									
Y2	0.333	0.667								
Y3	0.143	0.286	0.571							
Y4	0.067	0.133	0.267	0.533						
Y5	0.032	0.065	0.129	0.258	0.516					
Y6	0.016	0.032	0.063	0.127	0.254	0.508				
Y7	0.008	0.016	0.031	0.063	0.126	0.252	0.504			
Y8	0.004	0.008	0.016	0.031	0.063	0.125	0.251	0.502		
Y9	0.002	0.004	0.008	0.016	0.031	0.063	0.125	0.250	0.501	
Y10	0.001	0.002	0.004	0.008	0.016	0.031	0.063	0.125	0.250	0.500

RATIO METHOD WEIGHTS ASSUMING POPULATIONS HALVED EACH YEAR

YEAR	OBS 1	OBS 2	OBS 3	OBS 4	OBS 5	OBS 6	OBS 7	OBS 8	OBS 9	OBS 10
Y1	1.000									
Y2	0.667	0.333								
Y3	0.571	0.286	0.143							
Y4	0.533	0.267	0.133	0.067						
Y5	0.516	0.258	0.129	0.065	0.032					
Y6	0.508	0.254	0.127	0.063	0.032	0.016				
Y7	0.504	0.252	0.126	0.063	0.031	0.016	0.008			
Y8	0.502	0.251	0.125	0.063	0.031	0.016	0.008	0.004		
Y9	0.501	0.250	0.125	0.063	0.031	0.016	0.008	0.004	0.002	
Y10	0.500	0.250	0.125	0.063	0.031	0.016	0.008	0.004	0.002	0.001

In summary, with Exponential Smoothing, the user controls the degree of stability and adaptability of the forecast and the influence of the TRF through the smoothing weight. A large weight produces less stable more responsive forecasts. Lowering the weight increases stability and decreases responsiveness. Distorted values have a greater initial impact on the forecast with a higher weight than a lower weight, but they are also filtered out of the forecast quicker with a higher weight than a lower weight. With the Ratio Method, the degree of stability and adaptability depends on changes in population size; therefore, it is beyond the control of the user and is unknown. Population increases create more responsive, less stable forecasts. Population decreases produce more stable, less responsive forecasts.

2. <u>BRF Volatility and Trend</u>. We defined an item's BRF as volatile if it had a percentage change of greater than 25% between any two consecutive years. Only one change between any two consecutive years was necessary to be defined as volatile. We used Pearson's correlation coefficients (r) (see APPENDIX D for r computation) to measure trend over the entire 10 year BRF history file (1,269,864 items, including some obsolete items).

Based on our criteria, 69% of the items had volatile BRFs. We found that most of the changes were decreases. The largest group (13% of the volatile items) had only one decrease and eight years of constant BRFs. The next largest groups (6%, 4%, and 3% of the volatile items) had nine straight years of BRF decreases, eight years of decreases and one BRF increase, and eight years of decreases and one year of no change, respectively.

Items with stable BRFs had little or no history. Of the items with stable BRFs, 43% had no BRFs for the first nine years; i.e., they had no chance to be volatile. We used statistical t-tests and F tests (see APPENDIX D) to determine that the volatile items had significantly larger average 3M populations than the stable items. As shown in TABLE III, the 3M populations vary more and are larger for the volatile items than the stable items. This implies that the volatile items have more opportunity for demand history.

TABLE III

3M POPULATION

	STABLE ITEMS	VOLATILE ITEMS
AVERAGE	140.0	409.2
STD DEVIATION	8,973.4	13,936.3

In the previous section, we found that the current policy places a large influence on the TRF. Therefore, we recomputed the BRF using the data on the history file with Exponential Smoothing but without TRFs, manual changes, ordnance freezes, and overrides; i.e., the subjective portion of the forecast was eliminated. The smoothing weight was held to .4 (current policy) and the first observed usage rate was the starting point. (NOTE: Prior to 1985, if a usage rate was greater than the BRF, a weight of .8 was used. Since this policy has been changed, and most of the volatile items had decreases rather than increases, we chose to use the current policy (α = .4) over the entire time period.) TABLE IV shows that eliminating the subjective portion of the forecast produced about 2.5 times more items with stable BRFs. This proves that the TRFs, manual changes, ordnance freezes, and overrides cause most of the BRF unstability, not Exponential Smoothing.

TABLE IV

BRF VOLATILITY

	SYSTEM BRF	RECOMPUTED BRF (W/O SUBJECTIVES)
VOLATILE	878,246	310,667
STABLE	391,618	959,197

We computed Pearson's correlation coefficients (r) to measure trend over time for the System BRFs, 3M usage rates (3M demand divided by 3M population), and Recomputed BRFs without subjectives. We used this statistic to determine whether the actual 3M rates are trending, and whether the current System BRF and/or the Recomputed BRFs show the same trend pattern.

The correlation coefficients, rounded to the nearest tenth, for the System BRF, the 3M usage rate, and Recomputed BRF without subjectives are displayed in TABLE V. We also provide the number and percent of items associated with each value of r. A value of 0 indicates no trend. Viewing TABLE V, we see that System BRFs (current method as found on the 10 year history file) are more likely to have a trend than 3M usage rates. The correlation coefficients were zero for only 7% of the System BRFs as compared to 78% for 3M usage rates. For the TABLE V items, Pearson's coefficient tells us we can be over 95% confident that items with an r greater than .5 or less than -.5 are trending. Thus we are over 95% confident that 57% of the item System BRFs have trend (33% negative trend and 24% positive trend) compared to 4% of the item 3M usage rates (2% negative trend and 2% positive trend). Eliminating the subjective portion of the forecast yields forecasts which have correlation coefficients closer to 3M usage rates; that is, the Recomputed BRFs trend closer to actual 3M usage rates than the System BRFs. Now, 73% of the item forecasts have values of zero for r. Only 15% of the item forecasts have a trend (5% negative trend and 10% positive trend).

TABLE V

TREND ANALYSIS

r	SYSTEM BRF		3M USAGE RATE		RECOMPUTED BRFs W/O SUBJECTIVES	
	NUMBER	PERCENT	NUMBER	PERCENT	NUMBER	PERCENT
- 1	13,073	1	32	0	2,838	0
-.9	169,307	13	926	0	29,762	2
-.8	137,383	11	3,313	0	12,804	1
-.7	61,708	5	8,246	1	8,189	1
-.6	36,811	3	9,516	1	16,102	1
-.5	97,761	8	22,027	2	10,926	1
-.4	18,984	2	20,604	2	7,823	1
-.3	31,583	3	20,222	2	14,720	1
-.2	23,122	2	17,874	1	7,499	1
-.1	23,354	2	19,571	2	12,676	1
0	90,958	7	990,562	78	923,459	73
.1	23,036	2	22,436	2	8,771	1
.2	25,851	2	21,464	2	16,210	1
.3	44,540	4	23,529	2	10,554	1
.4	31,376	3	25,295	2	19,819	2
.5	132,490	10	34,115	3	40,319	3
.6	29,052	2	12,531	1	41,659	3
.7	75,444	6	10,819	1	44,332	4
.8	113,550	9	5,215	0	27,374	2
.9	88,291	7	1,516	0	12,780	1
1	180	0	41	0	1,238	0

In summary, examining the volatility and trend of System BRFs, we found that items with volatile BRFs tend to have a small number of changes (which are decreasing) and items with stable BRFs have little to no history. The System BRF trends more than the actual 3M usage rate; i.e., 57% of the item System BRFs show trend versus 4% for item 3M usage rates. Recomputing BRFs from the data without subjectives (i.e., eliminating TRFs, manual changes, ordnance freezes, and overrides) produces a more stable forecast and one that trends closer to actual 3M usage rates. The number of items with stable BRFs increased by about 150% and the items with trend decreased from 57% to 15%. This proves that volatility and trend are largely due to TRFs, manual changes, ordnance freezes, and overrides, not to Exponential Smoothing or alternative data sources.

3. CASREP Requisition/Problem Item Analysis. We reviewed sample CASREP "problem" items to determine whether the BRF was a major cause of the CASREP. Most CASREP requisitions only occurred once in the two year period. Of the 43,024 items with CASREP requisitions, 71% only occurred once. Based on the criteria of the previous section, 59% of the CASREP items had volatile System BRFs (64% of these had only one CASREP requisition). Out of the 57 ship types reporting CASREP requisitions, 54% of the CASREPs were from five ship types; i.e., SSN, DD, DDG, FF, and FFG.

Using Pearson's correlation coefficients from the previous section ($r > |.5|$), we investigated whether or not there was a problem with System BRFs and 3M usage rates trending in opposite directions. This was not a problem. Of the 1,001 items with BRFs having positive trends and 3M usage rates with negative trends, 12% had CASREPs (57% of these had only one CASREP requisition). The number of CASREP requisitions ranged from zero to 17. We found 5,180 items having System BRFs with negative trends and 3M usage rates with positive trends. Only 14% of these items had CASREP requisitions (62% of these occurred only once). The number of these requisitions ranged from zero to 36. For most of these 5,180 items, the TRF is larger than the 3M usage rate, and over time the System BRF and 3M usage rates are trending toward each other.

For analysis purposes, we defined a problem item as an item with more than 50 CASREPs in the two year period. Based on our criteria, 26 items were identified as problem items. We readily concede that items with 50 or less CASREPs may also be problem items; however, limiting the sample to a small number of items allowed us to do more in-depth analysis.

Some of the problems we found with these items were not related to the BRF. First, 3M population data was not always available for these items on the 10 year history file, even during the period when the CASREPs occurred. Second, these items' BRFs tended to be large. The largest, smallest, most recent, and oldest BRFs for the NIINs of these items are displayed in TABLE VI, along with the number of years of BRF history (an asterisk denotes items which had gaps where the history was missing or unavailable). For 62% of these items, the smallest BRF was at least .1, and for 69% of the problem items, the most recent BRF was at least .1. Under MOD-FLSIP rules, an item with an expected demand (BRF * Population) of at least .1 and an Item Mission Essentiality Code (IMEC) greater than 2 has at least one Minimum Replacement Unit (MRU) stocked. Since these items had more than 50 CASREPs, we can assume that their IMECs were 3 or more. Therefore, even with a population as low as one, we can assume at least one MRU was stocked for most of these items. This implies that the problems associated with these items may be related to the population data, or the model (not enough depth), or supply availability (assets not available in the system). Finally, all of the problem items are repairables. Repairable items (even appearing multiple times in an equipment) must be CASREPed separately. Therefore, one equipment failure can generate multiple CASREPs for the same item, as well as different items.

TABLE VI

PROBLEM ITEMS' BRFs

NIIN	MIN BRF	MAX BRF	MOST RECENT	OLDEST	NUMBER OF YEARS BRF HIST
001221493	0.350	0.490	0.380	0.350	9*
003635619	0.260	0.970	0.970	0.260	10
004662258	1.000	1.700	1.300	1.000	10
006137235	0.016	0.110	0.020	0.110	10
008690109	0.003	0.007	0.003	0.005	10
009495591	0.180	0.400	0.300	0.190	10
010224765	0.020	0.180	0.038	0.020	10
010498099	0.230	0.600	0.230	0.600	10
010543301	0.110	0.970	0.110	0.500	10
010673688	0.100	0.180	0.160	0.100	10
010804324	0.120	0.800	0.120	0.800	9
010822927	0.066	0.160	0.160	0.100	10
010898895	0.045	0.300	0.093	0.300	10
010965933	0.500	0.780	0.670	0.500	6
011041308	0.240	0.431	0.300	0.431	9*
011099480	2.000	2.100	2.100	2.000	10
011187279	0.038	0.120	0.110	0.120	10
011442593	0.100	0.270	0.250	0.100	6
011513058	0.170	0.490	0.490	0.170	5
011575880	0.001	0.780	0.001	0.780	8
011577009	0.046	0.164	0.062	0.046	9
011586889	0.100	0.650	0.100	0.650	5*
011637619	0.034	0.170	0.034	0.170	9
011769727	0.001	0.330	0.001	0.330	8
011979852	0.497	0.497	0.497	0.497	4*
012126298	0.309	0.309	0.309	0.309	3*

*Denotes history not consecutive

We found inconsistent usage rates for these items, both over time and between 3M, MLSF, and system data. The autocorrelation coefficients (this is another way to measure trend which requires a visual confirmation) for all three data source usage rates dropped to zero after one time lag. This indicates the data is stationary (a flat nonsloping trend). Therefore, either the Exponential Smoothing method or the Ratio Method is appropriate, as neither of these methods allow for a linear trend.

Most of the items have significant BRF histories (see TABLE VI). For some items, the BRF history is not consistent. However, 62% of the problem items have nine or more consecutive years of history. For these items, the TRF should account for no more than 1.7% of the forecast (see TABLE I). Therefore, the TRF should not be a factor in these BRFs.

Although not shown in this report, we plotted the System BRF, the Recomputed BRF, the Ratio Method BRF, and the usage rates over time. The system forecast was the actual BRF value on the 10 year history file. In most instances, the System BRF exceeded the usage rates. The System BRF did not always follow an item's usage rate pattern. For some items, a BRF change was exaggerated as compared to the actual usage rate change. This indicates manual manipulation of BRFs for some reason.

We next analyzed whether or not the Ratio Method provided a more stable forecast than Exponential Smoothing. A previous study (reference 1 to APPENDIX A) compared the current method and the Ratio Method with respect to accuracy and COSAL effectiveness. However, it did not distinguish differences in the forecasts due to TRF transitioning. We wanted to compare method to method with no TRF transitioning effect. The Recomputed BRF without subjectives allows us to compare pure Exponential Smoothing with the Ratio Method. The Ratio Method BRF was computed using 3M data only (as per APPENDIX B), but no TRF transition was made. Therefore, if an item had 3M data only, the Recomputed BRF and the Ratio Method BRF would use the same data.

Looking at the forecasts over time, both the Ratio Method and Exponential Smoothing BRFs varied, but not more than the actual usage rates. To determine if one forecast varied more than the other, we conducted an F test on their variances for each item (see APPENDIX D). The F test indicated that the Exponential Smoothing BRF varied more than the Ratio BRF (at the 5% level of significance) for five of the 26 items. For two items, this was a matter of using alternative data versus 3M data only forecasts. One of these two items never had any 3M data (other sources were available); i.e., there was no ratio forecast (in reality it would be the TRF). The other item had 3M data for only the first two years (1978 and 1979); therefore, the forecast was not updated by the Ratio Method, but was with Exponential Smoothing. The other three items contained extreme observations that should have qualified for a manual review. Ratio Method did not vary more than the Exponential Smoothing for any of the 26 items.

In summary, we found CASREPed requisitions occurred for a relatively small number of items (43,024) and most occurred only once in the two year history (71%). Items with BRF and 3M usage rates with opposite trends were not a factor. Twenty-six items were identified as problem items for in-depth analysis. Population data (3M) was not always available for these items on the 10 year history file. Most of these items had BRFs large enough to ensure stockage, regardless of population size, and the BRF was often larger than the observed usage rates. Since most of these items have significant BRF histories, the TRF should have been filtered out of the forecast and was not considered a problem. Usage rates for the problem items were found to be inconsistent and stationary. Evidence of manual BRF changes were observed. Exponential Smoothing with a weight of .4 tends to not vary more than the Ratio Method.

4. The PHALANX Data. Here we narrowed the scope of our analysis by targeting an equipment. The PHALANX was selected because it is a high profile problem equipment with a reasonable number of items (3,336). We were able to match 82% of the items with the 10 year history file. First, we examined the CASREP requisitions. Then we compared Exponential Smoothing and the Ratio

Method. Keep in mind, our purpose here was to gain insight, not to determine the best method.

The number of items, percent of items with no CASREP requisitions, percent of items with one CASREP requisition, and the maximum number of CASREP requisitons are shown in TABLE VII for the overall equipment and grouped by Cog. The majority of items did not have chronic problems. Only 12% had a CASREP requisition in the two year period. Multiple CASREPs occurred for only 5% of the items. If an item had a problem, it probably was a 7 Cog (repairable item). For 7 cog items, 18% had more than one CASREP requisition in the two year period, compared to 4% and 2% for 1 and 9 Cog items. The number of CASREP requisitions for an item also ranged higher for 7 Cog than for 1 and 9 Cog. For the PHALANX, 7 Cog items represented a small number of items (12%), but a large proportion of CASREP requisitions (29%) and multiple CASREPs (43%).

TABLE VII

PHALANX CASREP ANALYSIS

	NO. ITEMS	PERCENT NO CASREPS	PERCENT ONE CASREP	MAX CASREPs
OVERALL	3,336	88	7	73
1 COG	711	88	8	12
7 COG	396	71	11	73
9 COG	2,220	92	6	16

Since the forecast performance measurements (stability, ME, and MSE) are only useful when comparing methods (we cannot say a method is stable or accurate; we can only say a method is more stable or more accurate than another method), we recomputed BRFs from the data using Exponential Smoothing and the Ratio Method (the methods available under Resystemization). Again, our interest is in comparing methods. Therefore, we ignored the System BRF (which has manual changes, ordnance freezes, overrides, and TRFs) and TRF transitioning. For Exponential Smoothing, we used the two weights from our first discussion comparing Exponential Smoothing and the Ratio Method; i.e., .4 and .1. As computed previously, the first observed usage rate was the starting point and the forecast was updated even when 3M data was unavailable. The Ratio forecast was computed as in the previous section with 3M data only.

Since we have a large number of items, we are concerned about the general tendency of the performance measurements. Therefore, we will use the mean and median to measure central tendency and the standard deviation to measure the dispersion. The mean is the arithmetic average. Each observation has the same weight. Therefore, it is sensitive to a few extreme values. The median is simply the midpoint. Half the observations are larger and the other half are smaller than the median. Extreme values have no influence on the median. The standard deviation measures how items vary around the mean. If all the observations are close to the mean, then the standard deviation will be small; i.e., there is little dispersion. We calculated the mean, median, and standard deviations for the ME and MSE for the two forecast methods, along with the smallest and largest errors.

We first analyzed the ME and MSE for the entire PHALANX data set (TABLE VIII). To eliminate the effect of using alternative data when 3M data is unavailable, we then performed the same analysis for the subset of PHALANX items containing 10 years of 3M data (TABLE IX). Finally, to determine the impact of updating with alternative data sources, we again performed the same analysis with the entire PHALANX data set, but we did not update the forecast when 3M data was unavailable (TABLE X).

Interpreting TABLE VIII, we see that in terms of ME, Exponential Smoothing with a weight of .4 is best, Exponential Smoothing with a weight of .1 is worst, and the Ratio Method is in between. We also note that the Ratio Method tends to underforecast (positive mean ME), while the other methods tend to overforecast (negative mean ME). The three methods are close in terms of the MSE.

TABLE VIII

PHALANX DATA SET - ALL THE DATA
3,336 ITEMS

	EXPONENTIAL SMOOTHING				RATIO METHOD	
	WEIGHT = .4		WEIGHT = .1			
	ME	MSE	ME	MSE	ME	MSE
MEAN	-0.289	22,956.4	-0.306	19,782.3	0.211	20,838.8
STD DEV	13.238	1,302,679	40.366	1,079,588	17.056	1,189,444
MAXIMUM	125.641	75,230,859	1439.87	62,259,066	935.627	68,695,920
MEDIAN	0	.000008	0	.000009	0	.000007
MINIMUM	-727.2	0	-1799.09	0	-298.83	0

Since the Ratio Method does not update the forecast when 3M data is unavailable and Exponential Smoothing does, we still do not have a true comparison of methods. We need to eliminate the effect of using alternative data when 3M data is unavailable. TABLE IX contains only PHALANX items with 10 years of 3M data. Now, all the forecasts use the same data, and we can obtain a true comparison of methods. Interpreting TABLE IX, we see that for the ME, the Ratio Method is the best, while Exponential Smoothing with a weight of .1 is worst. The same results hold true for the MSE.

TABLE IX

PHALANX DATA SET - 10 YEARS OF 3M POP
883 ITEMS

	EXPONENTIAL SMOOTHING				RATIO METHOD	
	WEIGHT = .4		WEIGHT = .1			
	ME	MSE	ME	MSE	ME	MSE
MEAN	-0.9	1408.6	-2.25	3971.4	-0.337	912.1
STD DEV	24.5	40,781	60.6	116,812	10.2	26,121
MAXIMUM	1.9	1,211,629	10.7	3,471,071	41.6	775,917
MEDIAN	-0.00004	0.00057	-0.0002	0.0008	0	0.0006
MINIMUM	-727.2	0	-1799.09	0	-298.83	0

Since these items had 10 years of 3M data, they are a subset of the entire PHALANX data set. That is, nothing different was done between TABLEs VIII and IX. We just isolated those items with 10 years of 3M data. These items may not give us a clear picture of the impact of updating/not updating with alternative data sources. Therefore, we again recomputed the Exponential Smoothing with a weight of .4 for the entire PHALANX data set and used only 3M data; i.e., we did not update the forecast when 3M data was unavailable. (We only recomputed the .4 forecasts because we wanted a concise comparison and .4 is the current policy.)

TABLE X contains the comparisons of Exponential Smoothing (weight of .4) using alternative data (benchmark) and 3M data only (alternative). Comparing the two Exponential Smoothing alternatives, we see that for both the ME and MSE, the method using alternative data sources is better than the 3M data only alternative. Overall, the evidence in TABLE X indicates that accuracy is lost by not updating the forecast when 3M data is unavailable. This loss of accuracy results in larger overforecasts.

TABLE X

PHALANX DATA SET - ALL THE DATA
3,336 ITEMS
EXPONENTIAL SMOOTHING WEIGHT = .4

	EXPONENTIAL SMOOTHING			
	BENCHMARK		ALTERNATIVE	
	ME	MSE	ME	MSE
MEAN	-0.289	22,956.4	-0.458	25,885.5
STD DEV	13.238	1,302,679	18.0422	1,473,485
MAXIMUM	125.641	75,230,859	25.14	85,097,483
MEDIAN	0	.000008	0	.000006
MINIMUM	-727.2	0	-745.314	0

NOTE: BENCHMARK = IF NO 3M POP EXISTS, THEN UPDATE FORECAST USING ALTERNATIVE DATA SOURCES
ALTERNATIVE = IF NO 3M POP EXISTS, THEN NO FORECAST UPDATE IS MADE

Now, we will look at stability. Here we examined the fast moving items; i.e., items with 10 years of positive 3M usage rates. We selected these 344 fast moving items for several reasons. As discussed previously, we used confidence limits on 3M data; therefore, items with no 3M data would be eliminated. Since the methods use the same data, we obtain a better comparison. Finally, the fast moving items should be the most variable. Exponential Smoothing with a weight of .1 was the least likely method to have an item's forecast exceed the limits (51% of the forecasts were within the confidence limits all 10 years); while, Exponential Smoothing with a weight of .4 was the most likely method to have a forecast exceed the limits (only 10% of the forecasts were within the confidence limits all 10 years). The Ratio Method was in between (36% of the forecasts were within the limits all 10 years). However, once an item's forecast exceeded the limits, it was most likely to stay outside using Exponential Smoothing with a weight of .1. It was least likely to stay outside using Exponential Smoothing with a weight of .4. Both the Ratio Method and Exponential Smoothing with a weight of .1 had item forecasts exceed the limits all 10 years. Exponential Smoothing with a .4 weight had no item forecast exceed the limits for more than six years. Exponential Smoothing with a weight of .4 had a standard deviation of 1.3. With a weight of .1, the measure of dispersion increased to 3.8. The Ratio Method was in between.

In summary, with Exponential Smoothing, lowering the weight increases the stability but reduces the adaptability of the forecast. Extreme values will have less impact on the forecast initially with a lower weight than with a higher weight; but over time, extreme values have a larger influence with the lower weight than the higher weight. This confirms the theoretical analysis presented earlier. Viewing TABLEs VIII and IX, we see the extreme values (the minimum and maximum ME) are larger for the lower weight than the higher weight. We have also seen that forecasts are more stable with a lower weight, but once they become distorted are less likely to self correct. With the higher weight the forecasts are more likely to become distorted, but are also quicker to readjust. The Ratio Method is somewhere in between.

With respect to CASREPs, the PHALANX data findings were consistent. Relatively few PHALANX items had CASREPs, and for those that did, most only occurred once in the two year period. Only 5% of the PHALANX items had more than one CASREP. Furthermore, if an item has a problem it is most likely a repairable item. Repairable items (7 Cog) represent a small percent of the PHALANX items (12%), but a large proportion of the CASREP requisitions (29% of all CASREP requisitions and 43% of multiple CASREPs).

Comparing forecast methods, Exponential Smoothing has a tendency to overforecast and the Ratio Method tends to underforecast when considering all items. Accuracy may be lost by not updating forecasts when 3M data is unavailable.

5. Conclusions - BRF Problem Analysis. With Exponential Smoothing, forecasts tend to err on the high side (overforecast). For the Ratio Method, forecasts tend to err on the low side (underforecast). Eliminating the subjective portion of the forecast (TRF, manual changes, ordnance freezes, and overrides) and using a constant smoothing weight of .4, we obtained BRFs that trend closer to actual 3M usage rates than current BRFs and are more stable. These forecasts vary no more than the Ratio Method forecasts. Lowering the smoothing weight to .1 decreases the magnitude of overall errors, but the magnitude of the extreme errors increased. The forecast was less likely to exceed 95% confidence limits for average 3M usage rates than with a weight of .4 or the Ratio Method. However, once the limits were exceeded, the forecasts were more apt to self adjust/correct with the higher weight (.4) than with the Ratio Method or the lower weight (.1). 3M data forecasts may not be as accurate as forecasts which also use other data when 3M is unavailable. These errors were generally larger overforecasts.

Item support problems may not be related to the BRF but to data, depth (the model), and/or supply availability. CASREP requisitions occurred for a relatively small number of items and most (71%) only occurred once in a two year period. Only 5% of the PHALANX data had more than one CASREP requisition. Most of the CASREP items analyzed had BRFs large enough to ensure shipboard stockage, and the BRF was often larger than the observed usage rate.

B. BRF IMPROVEMENT.

From the previous section we found:

- BRF instability and trend primarily result from the subjective portion of the forecast - TRF, manual changes, ordnance freezes, and overrides.
- Short term extreme data observations also produce excessive variation.
- Based on the number of items with zero correlation coefficients (and the autocorrelation coefficients for the problem items), most items are stationary. Therefore, single Exponential Smoothing is an appropriate forecast method.
- 3M data is not always available to update the BRF.

Therefore, we need to:

- Examine alternatives for transitioning the TRF into the forecast.
- Minimize the impact of short term extreme data observations.
- Determine the impact of updating/not updating the forecast when 3M data is unavailable.

1. Alternative Forecast Methods. The following discussion gives a brief overview of the forecasting methods considered as alternatives to the current method. For a detailed description and the mathematical formulas see APPENDIX B.

The DDG 52 RBS working group is using (and proposing) a Bayesian method. Bayesian forecast methods use Bayes' theorem to update a subjective forecast (TRF) with objective data (usage rate). The weights are similar to the Ratio Method (see APPENDIX C). The tradeoff between adaptability and stability depends on changes in population size; i.e., increases in population size produce a more adaptable forecast, and decreases produce a more stable one. The amount of influence given the subjective portion versus the objective portion of the forecast depends on TRF size. For small values, the TRF has more influence than the objective data. With large values, the objective data has more influence than the TRF. With the Bayes method, the weight is an unknown variable beyond the control of the user.

We compute the Bayes forecast two ways. One forecast uses 3M data only. The other uses alternative data sources to update the forecast when 3M data is unavailable. The priority of data selection is the same as under current policy; i.e., 3M, then MLSF, then system demand.

Since the Ratio Method is included in Resystemization, we also reevaluate it. This method is computed as discussed previously and described in APPENDIX B. (Currently, the Ratio Method is defined to only use 3M data.)

We next develop an alternative specifically designed to address the problems identified earlier in the report. As shown previously, one of the problems with the current procedure is related to TRF transitioning. The forecast can also be improved by eliminating excessive short term data observations. If a lower weight is used, accuracy and stability improve for some items. For other items, they get worse. The same can be said for using 3M data only versus updating the forecast with alternative data when 3M is unavailable.

The smoothing weight and data concerns can be handled easily. We can test which forecast performs better for the majority of the items. Therefore, we compute the exponential forecasts with 3M data only and using alternative data when 3M is unavailable. A high weight and a low weight will be used. We chose .4 and .2 as the smoothing weights, because under current policy, .4 is used at the consumer level and 2 at the wholesale level.

TRF transitioning and minimizing excessive short term data changes are more difficult problems. We cannot discard the TRF. For some time period, the TRF is the only estimate we have of the usage rate. Also, Exponential Smoothing must have observed usage to identify a pattern (an average or level) and project that pattern into the future. Furthermore, we need to have observed some usage rates to determine what is an excessive data change.

As an alternative to current procedures, we adjust the TRF over the development period with a probability distribution. We can treat the TRF as a point estimate. The question becomes: given this estimate of the usage rate (TRF), is the usage rate observed realistic? If it is, then the TRF is a reasonable estimate and we do not adjust it. Otherwise, we adjust the TRF to a more reasonable level. We selected the exponential probability distribution since it is often used to model failure rates, has been shown to be an appropriate distribution to model demand (reference 2 of APPENDIX A), and requires only one estimated parameter. For this study, the development period was three years. We set a lower limit based on what we would expect to observe 50% of the time. We set an upper limit based on what we expect to observe 95% of the time. For the first three years (forecasts 2 through 4), we adjust the TRF if the usage rates observed are larger or smaller than we would expect, given that estimate of the usage rate. In the fourth year, we switch to Exponential Smoothing, if we have observed four years of data. If we have not, then we continue to adjust the TRF until we have observed four consecutive years of usage data. The intention here is to develop a pattern. The development period should be long enough to develop a pattern.

With Exponential Smoothing, the starting point can be critical; therefore, we considered two different starting points. First, we used the adjusted TRF described above. Since the TRF has been adjusted over the development period, it may now provide a valid starting point.

For the other starting point, we used an average of the first three years of observed usage rates. A pattern should now be established and we no longer need the TRF. However, a given usage rate observation may be distorted. By using an average, we do not put too much weight on one observation; i.e., we want the distortion to filter out quickly. Starting in the next year (the second year after switching to Exponential Smoothing), we "windsorize" the most recent observation using the probability distribution limits previously used to adjust the TRF. Here, we substitute the smoothed forecast in place of the TRF. That is, based on what we have observed in the past, we identify extreme observations for the usage rate as those that fall outside upper and lower limits. (This method was developed by Tukey; see reference 3 of APPENDIX A.) If an observation is outside the limits, we will adjust the observation to the lower or upper limit. The adjusted usage rates will then be smoothed into the forecast. Thus, excessive short term data changes will be minimized.

For example, a BRF in year 7 with a value of 1.0, would cause the 8^{th} year's observed usage rate to be constrained between .7 and 3.0. Given this constraint on the usage rate, the BRF for year 8 would be restricted to a value between .88 and 1.8 in year 8. We want to find or establish the pattern of the data (the usage rates). Short term changes distort the forecast away from the true pattern of the data. Therefore, we want to ignore them. However, we can not distinguish between a short term change, the pattern, or a new emerging pattern until after the fact.

To illustrate this procedure, assume an item has the usage and population data for an 11 year period shown in TABLE XI and a TRF of 2. Here we chose a smoothing weight of .4 and the average data starting point method.

TABLE XI

SAMPLE ALTERNATIVE BRF COMPUTATION

YEAR	USAGE	POPULATION	BRF W/AVG DATA STARTING POINT
1	10	1	2
2	12	2	6
3	10	2	6
4	9	3	6
5	15	5	5.4
6	100	2	4.8
7	10	2	8.6
8	11	1	7.6
9	12	2	9.0
10	10	5	7.9
11	9	3	7.0

The BRFs for the first four years are the TRFs adjusted over the development period with the exponential probability distribution. Since we have four consecutive years of observed data, the remaining BRFs are computed using exponential smoothing, but with the most recent observation "windsorized" to minimize short term data changes using the probability distribution limits previously used to adjust the TRF. For example, the usage and population data in year 6 are used to compute the BRF shown in year 7. The usage rate of 50 (100/2) for year 6 is excessive compared to the previous years' usage rates. Thus, we "windsorize" the data by setting the year 6 usage rate to 14.4, or three times the BRF (estimated usage rate for year 6). Then the 14.4 is smoothed with the forecast (4.8), using a smoothing weight of .4 to get the final result of 8.6 for year 7.

We have now identified three different forecast methods with various parameters and options, for a total of 11 alternatives to be evaluated.

- Bayes using alternative data sources when 3M is unavailable.
- Bayes using 3M data only.
- Ratio Method using 3M data only.
- Exponential Smoothing using alternative data sources when 3M is unavailable.
  - $\alpha = .2$ average data starting point
  - $\alpha = .4$ average data starting point
  - $\alpha = .2$ TRF starting point
  - $\alpha = .4$ TRF starting point
- Exponential Smoothing using 3M data only.
  - $\alpha = .2$ average data starting point
  - $\alpha = .4$ average data starting point
  - $\alpha = .2$ TRF starting point
  - $\alpha = .4$ TRF starting point

2. Selecting the Best Alternative. To begin our analysis, we first need to determine the TRF. Therefore, we eliminated (from the 10 year history file) items with BRFs in the oldest year, then we can assume that the first observed BRF is the item's TRF. To obtain meaningful measurements of forecast performance, items also had to have at least five years of history, with at least three years of the history containing 3M data. This left 128,215 items to measure forecast performance; of which, 91,543 were identified as readiness contributors by their FSC codes as specified by NSLC.

Since there are a large number of items, forecast performance was evaluated based on criteria which measures the tendency of these items for a particular alternative. Forecast performance was ranked based on that criterion from one for best performance to 11 for worst performance. Methods which tie for a particular performance element were assigned the average of the ranks; e.g., two methods which tie for best performance would each be given a rank of 1.5 (1+2/2). The ranks were totaled across the three performance measurements: stability, accuracy with respect to bias and extremes, and magnitude of the errors. For each performance measurement, we used five different statistics. Therefore, no one performance measurement dominated the selection process. Since each statistic measured the performance differently, we have a fair representation of how well a forecast method performs. We considered the method with the lowest rank sum to be the best alternative.

To determine how the alternative methods compare to the current procedure, we computed the BRF using the current method and same data (TRFs and usage rates) as the alternatives, with no manual changes, ordnance freezes, or overrides. This is the Benchmark; however, we did not consider the Benchmark in the ranking. Ranks shown for the Benchmark reflect its performance relative to the performance of the 11 alternatives. The Benchmark ranks are provided only to facilitate comparisons; i.e., to determine whether or not the alternatives provide better results than the current procedure. For example, the Benchmark's rank was set equal to an alternative's rank when their numbers (mean, median, etc.) were the same.

a. Stability. We defined stability based on a method's ability to forecast within the 95% confidence limits of the average 3M usage rate. Therefore, we used the number of items always within these limits as a performance element. The method with the most items has the best performance, the least items has the worst performance. We also measured the mean number of times items were outside the limits. A forecast with a large mean has a higher number of items outside the limits for more years than one with a smaller mean. Therefore, we assigned the best (lowest) rank to the method with the smallest mean and the highest (worst) rank to the largest mean. The median is not influenced by large or small values. Therefore, a forecast with a smaller median than another method is more stable. The mode is the most frequent observation. We assigned the best rank to the alternative with the largest number of items representing the smallest mode. Finally, we considered the forecast alternative with the smallest standard deviation to be the most stable, since it is not as widely dispersed as the other methods (has more values closer to the mean).

The stability results are displayed for the readiness contributors in TABLE XII and for all items in TABLE XIII. The results are consistent for both data sets. The lowest cumulative ranks were achieved by the Exponential Smoothing methods using alternative data sources and a data starting point with either weight. These two alternatives were also ranked top in each category with the exception of the **"NO. ITEMS ALWAYS W/I LIMITS"** category. But even there, the numerical differences were small. However, the overall rank difference between these two most stable alternatives and the methods available under Resystemization (Ratio and Benchmark) are large.

Forecasts using alternative data when 3M data was unavailable produced more stable forecasts. For the Bayes forecasts, using alternative data resulted in a few more items always within the limits (648 for readiness contributors/821 for all items). This effect was more pronounced for the Exponential Smoothing forecasts than the Bayes. With Exponential Smoothing (average data starting point either weight), the difference was primarily in the mode. The value for the mode was the same, but over 2.5 times as many items had that value for the alternative data forecast vs the 3M data only forecast. Consequently, the means and standard deviations are less for the alternative data forecast; i.e., using only 3M data produced forecasts that were outside of the limits for more years than using alternative data to update the forecast when 3M was unavailable.

For Exponential Smoothing, the average data starting point produced more stable forecasts than the TRF starting point regardless of the data source or weight. The means, medians, modes, and standard deviations are higher for any of the TRF starting point forecasts than the data starting point forecasts. (Note the median and modes for the data starting point forecasts are three years. For all other methods, the median and modes are six years. Since we discarded the first forecast, we would have used the adjusted TRF for a minimum of three years.)

We used ranks in order to make an objective decision with multiple criteria. A problem with ranking is that small differences may be inflated and large differences may appear small. Consequently, care was taken that no one category or statistic would dominate the selection process. To insure that this occurred and that one rank did not have an undue influence, we reviewed each statistic, comparing the selected method with the alternatives. With respect to stability (TABLEs XII and XIII), no method performed better than Exponential Smoothing using alternative data sources and an average data starting point based on the mean, median, mode, and standard deviation. The Bayes and Ratio alternatives did forecast within the limits for more items over the entire forecast horizon than Exponential Smoothing. However, the maximum difference was less than 3% of the universe.

TABLE XII

STABILITY FINDINGS

READINESS CONTRIBUTORS ONLY (91,543)
BASED ON 95% LIMITS ON AVG 3M USAGE RATES

	NO. ITEMS ALWAYS W/I LIMITS	RANK	NO. OF YEARS FORECAST OUT OF 95% LIMIT										SUM OF RANKS
			MEAN	RANK	MEDIAN	RANK	MODE	%	RANK	STD	RANK		
BAYES ALT DATA	21,462	1	5.2	5	6	8	6	24	7	3.2	9		30
BAYES 3M ONLY	20,814	2	5.3	7.5	6	8	6	24	7	3.2	9		33.5
RATIO 3M ONLY	16,653	11	5.3	7.5	6	8	6	24	7	3.1	5.5		39
EXPONENTIAL SMOOTHING:													
ALT DATA:													
α = .2 DATA ST PT	19,150	4	2.5	1.5	3	2.5	3	70	1.5	1.6	1.5		11
α = .4 DATA ST PT	18,971	6	2.5	1.5	3	2.5	3	70	1.5	1.6	1.5		13
α = .2 TRF ST PT	19,432	3	5.4	10.5	6	8	6	25	5.5	3.2	9		36
α = .4 TRF ST PT	19,017	5	5.3	7.5	6	8	6	24	7	3.2	9		36.5
3M ONLY:													
α = .2 DATA ST PT	18,706	8	3.7	3.5	3	2.5	3	26	3.5	2.5	3.5		21
α = .4 DATA ST PT	18,597	10	3.7	3.5	3	2.5	3	26	3.5	2.5	3.5		23
α = .2 TRF ST PT	18,944	7	5.4	10.5	6	8	6	25	5.5	3.2	9		40
α = .4 TRF ST PT	18,617	9	5.3	7.5	6	8	6	24	7	3.1	5.5		37
*BENCHMARK	16,167	12	5.4	10.5	6	8	6	25	5.5	3.1	5.5		41.5

NOTE: BENCHMARK RANKS ONLY DENOTE RANKING RELATIVE TO ALTERNATIVE FORECASTS.

TABLE XIII

STABILITY FINDINGS

ALL ITEMS (128,215)
BASED ON 95% LIMITS ON AVG. 3M USAGE RATES

	NO. ITEMS ALWAYS W/I LIMITS	RANK	NO. OF YEARS FORECAST OUT OF 95% LIMIT										SUM OF RANKS
			MEAN	RANK	MEDIAN	RANK	MODE	%	RANK	STD	RANK		
BAYES ALT DATA	30,158	1	5.1	5.5	6	8	6	26	8	3.3	11		33.5
BAYES 3M ONLY	29,337	2	5.1	5.5	6	8	6	26	8	3.2	8.5		32
RATIO 3M ONLY	23,099	11	5.2	8	6	8	6	26	8	3.1	5.5		44.5
EXPONENTIAL SMOOTHING:													
ALT DATA:													
α = .2 DATA ST PT	26,654	4	2.5	1.5	3	2.5	3	68	1.5	2.5	2.5		12
α = .4 DATA ST PT	26,398	7	2.5	1.5	3	2.5	3	68	1.5	2.5	2.5		15
α = .2 TRF ST PT	27,023	3	5.3	10.5	6	8	6	26	8	3.2	8.5		38
α = .4 TRF ST PT	26,465	5	5.2	8	6	8	6	26	8	3.2	9.5		37.5
3M ONLY:													
α = .2 DATA ST PT	26,103	8	3.4	3	3	2.5	3	26	3.5	2.5	2.5		19.5
α = .4 DATA ST PT	25,926	10	3.6	4	3	2.5	3	26	3.5	2.5	2.5		22.5
α = .2 TRF ST PT	26,421	6	5.3	10.5	6	8	6	26	8	3.2	8.5		41
α = .4 TRF ST PT	25,973	9	5.2	8	6	8	6	26	8	3.1	5.5		38.5
*BENCHMARK	22,513	12	5.3	10.5	6	8	6	26	8	3.1	5.5		44

NOTE: BENCHMARK RANKS ONLY DENOTE RANKING RELATIVE TO ALTERNATIVE FORECASTS.

b. Bias and Extremes. Negative MEs denote overforecasts (less risk of stockout). Therefore, we assigned the method(s) with the least negative mean and median MEs (i.e, the least overforecast) the best (lowest) ranks and the method with the most positive mean and median MEs (i.e., the greatest underforecast) the worst (highest) ranks. All overforecasts were ranked better than the underforecasts. Since one extreme value can influence the mean and we are interested in how well the forecast methods handle extreme values, the method(s) with the smallest (in absolute terms) maximum and minimum were considered as performing the best. The standard deviation also denotes how well a method handles extreme values. A method with a smaller standard deviation has less extreme values than one with a larger standard deviation.

The results with respect to bias and extremes are displayed for the readiness contributors in TABLE XIV and all items in TABLE XV. Here, the results are inconsistent between the readiness contributors and all items. Based on readiness contributors, the Exponential Smoothing (.4 smoothing weight) using alternative data with either starting point are the best performers with cumulative ranks of about 16. The current BRF method is next at 19. Across all items, all the methods except smoothing with the TRF starting point are ranked fairly close.

The results were also inconsistent between readiness contributors and all items for 3M only vs alternative data forecasts. For all items, for both the Bayes and Exponential Smoothing forecasts (all weights and starting points), the 3M only data forecasts outperformed the alternative data forecasts. For the readiness contributors, the alternative data outperformed the 3M data only forecasts. The average data starting point performed better for all items; while there was little difference between the two starting points for the readiness contributors.

With respect to bias, the smoothing method (.4 weight) with alternative data sources and average data starting point method does have a tendency to err by overforecasting. However, it does not overforecast as much as other alternatives. The medians were the second best for both the readiness contributors and all items. Furthermore, the mean performed second best for the readiness contributors and third best for all items. In terms of handling extreme values, this method was ranked near the middle. Compared to the methods available under Resystemization, this method does not overforecast as much as the Benchmark and it does not handle extreme forecast errors as well. The Ratio Method consistently underforecasts based on the medians and means.

TABLE XIV

MEAN ERROR FINDINGS

READINESS BASED CONTRIBUTORS ONLY (96,543)

	MEAN	RANK	MEDIAN	RANK	MAX	RANK	MIN	RANK	STD	RANK	SUM OF RANK
BAYES ALT DATA	-.0012	1	-.0092	9	676.2	6	- 686.9	10.5	3.2	10.5	37
BAYES 3M ONLY	-.0075	4	-.0102	10	677.3	8	- 686.9	10.5	3.2	10.5	43
RATIO 3M ONLY	.0249	11	.0156	11	200.4	1	- 660.5	9	2.3	2.5	34.5
EXPONENTIAL SMOOTHING:											
ALT DATA:											
α = .2 DATA ST PT	-.0032	3	-.0046	1	659.6	4	- 509.1	7.5	2.7	8	23.5
α = .4 DATA ST PT	-.0020	2	-.0048	2	654.6	2	- 314.4	5.5	2.4	4.5	16
α = .2 TRF ST PT	.0050	9	-.0064	4	665.1	5	- 136.8	3.5	2.3	2.5	24
α = .4 TRF ST PT	.0019	7	-.0053	3	656.0	3	- 93.1	1.5	2.2	1	15.5
3M ONLY:											
α = .2 DATA ST PT	.0010	5	-.0076	5.5	676.9	7	- 509.1	7.5	2.9	9	34
α = .4 DATA ST PT	.0016	6	-.0076	5.5	677.8	9	- 314.4	5.5	2.6	7	33
α = .2 TRF ST PT	.0053	10	-.0086	8	679.1	10	- 136.8	3.5	2.5	6	37.5
α = .4 TRF ST PT	.0048	8	-.0080	7	679.4	11	- 93.1	1.5	2.4	4.5	32
*BENCHMARK	-.0089	5	-.0079	7	663.3	5	- 38.79	1	2.2	1	19

Highest rank assigned to:
Least Negative Mean
Least Negative Median
Smallest Max
Largest Min

<u>NOTE</u>: BENCHMARK RANKS ONLY DENOTE RANKING RELATIVE TO ALTERNATIVE FORECASTS.

TABLE XV

MEAN ERROR FINDINGS

ALL ITEMS (128,215)

	MEAN	RANK	MEDIAN	RANK	MAX	RANK	MIN	RANK	STD	RANK	SUM OF RANK
BAYES ALT DATA	.0134	4	-.0085	9	676.2	1	-1783.9	8	7.12	5	27
BAYES 3M ONLY	.0386	6	-.0096	10	677.3	2	- 686.9	5	4.63	1	24
RATIO 3M ONLY	.0010	2	.0149	11	1463.7	5	-1159.4	6	5.99	3	27
EXPONENTIAL SMOOTHING:											
ALT DATA:											
α = .2 DATA ST PT	-.0389	1	-.0044	1	903.9	3.5	-2816.7	11	11.06	11	27.5
α = .4 DATA ST PT	.0018	3	-.0046	2	1795.7	6.5	-2173.7	10	9.96	8	29.5
α = .2 TRF ST PT	.0467	8	-.0061	4	2889.8	10.5	-1587.2	7	10.63	10	39.5
α = .4 TRF ST PT	.0331	5	-.0051	3	2497.7	8.5	-1840.2	9	10.21	9	34.5
3M ONLY:											
α = .2 DATA ST PT	.0419	7	-.0072	5	903.9	3.5	- 575.7	4	4.98	2	21.5
α = .4 DATA ST PT	.0654	9	-.0073	6	1795.7	6.5	- 314.4	3	6.88	4	28.5
α = .2 TRF ST PT	.0871	11	-.0081	8	2889.8	10.5	- 149.3	2	9.4	7	38.5
α = .4 TRF ST PT	.0820	10	-.0076	7	2497.7	8.5	- 102.6	1	8.4	6	32.5
*BENCHMARK	-.008	2	-.0078	8	663.3	1	-1758.2	8	7.4	5	24

Highest rank assigned:
Least Negative Mean
Least Negative Median
Smallest Max
Largest Min
Smallest Std

NOTE: BENCHMARK RANKS ONLY DENOTE RANKING RELATIVE TO ALTERNATIVE FORECASTS.

c. Accuracy. Small MSEs indicate errors that are small in magnitude. Therefore, we considered the alternative(s) with the smallest mean and median MSEs to be the most accurate forecast. Since we have already considered extreme values, we now focus on the upper and lower quartiles. These statistics are similar to the median and insensitive to extreme values. The upper quartile has 25% of the observations greater than or equal and 75% less than or equal to it. The lower quartile has 75% of the observations greater than or equal and 25% of the observations less than or equal to it. They allow us to ignore one-fourth of the largest values and one-fourth of the smallest values. We assign the method with the smallest MSE quartile the best (lowest) rank. We use the standard deviation here also. Again, we consider the method with the smallest value for the standard deviation to be the best performer.

The results for accuracy with respect to the magnitude of the errors are displayed in TABLE XVI for the readiness contributors and TABLE XVII for all items. The MSEs will be small for the quartiles and medians. (This is the nature of the measurements. We would be concerned if they were not, given the small size of most BRFs and usage rates.)

The tables show that for readiness contributors and for all items, the Exponential Smoothing methods using alternative data and the TRF starting point (either weight), and 3M only with a weight of .4 and the TRF starting point are the best performers. For readiness contributors, the ranks were 15-17.5. The Exponential Smoothing method using alternative data sources, an average data starting point with a weight of .4 is close with a total rank of 22.

TABLE XVI

MEAN SQUARED ERROR FINDINGS

READINESS CONTRIBUTORS ONLY (96,543)

	MEAN	RANK	MEDIAN	RANK	UPPER QUARTILE	RANK	LOWER QUARTILE	RANK	STD	RANK	SUM OF RANK
BAYES ALT DATA	48.8	11	.0003	7.5	.0020	6	.00005	9	10105.9	11	44.5
BAYES 3M ONLY	48.5	10	.0004	10	.0029	10	.00006	10	10102.6	10	50
RATIO 3M ONLY	38.6	6	.0007	11	.0051	11	.00010	11	8814.8	7	46
EXPONENTIAL SMOOTHING:											
ALT DATA:											
α = .2 DATA ST PT	43.1	9	.0002	5	.0018	3.5	.00001	1.5	8911.6	8	27
α = .4 DATA ST PT	39.9	7	.0002	5	.0018	3.5	.00001	1.5	8696.1	5	22
α = .2 TRF ST PT	35.8	2	.0002	5	.0012	1	.00002	4	8601.8	3	15
α = .4 TRF ST PT	36.7	3.5	.0002	5	.0018	3.5	.00002	4	8601.7	1.5	17.5
3M ONLY:											
α = .2 DATA ST PT	42.4	8	.0003	7.5	.0028	8.5	.00003	7	8917.9	9	40
α = .4 DATA ST PT	38.5	5	.0003	7.5	.0027	7	.00003	7	8700.3	6	32.5
α = .2 TRF ST PT	35.7	1	.0003	7.5	.0028	8.5	.00003	7	8643.0	4	28
α = .4 TRF ST PT	36.7	3.5	.0002	5	.0018	3.5	.00002	4	8601.7	1.5	17.5
*BENCHMARK	34.4	1	.0003	7.5	.0029	10	.00003	7	8627.1	4	29.5

HIGHEST RANK TO THE SMALLEST VALUE

*NOTE: BENCHMARK RANKS ONLY DENOTE RANKING RELATIVE TO ALTERNATIVE FORECASTS

TABLE XVII

MEAN SQUARED ERROR FINDINGS

ALL ITEMS (128,215)

	MEAN	RANK	MEDIAN	RANK	UPPER QUARTILE	RANK	LOWER QUARTILE	RANK	STD	RANK	SUM OF RANK
BAYES ALT DATA	559.9	2	.0003	7.5	.0021	6	.00004	9	152,837	2	26.5
BAYES 3M ONLY	510.0	1	.0004	10	.0030	10	.00005	10	151,430	1	32
RATIO 3M ONLY	942.2	3	.0007	11	.0053	11	.00010	11	294,800	3	39
EXPONENTIAL SMOOTHING:											
ALT DATA:											
α = .2 DATA ST PT	1629.8	11	.0002	3	.0020	4	.00001	1.5	497,686	11	30.5
α = .4 DATA ST PT	1289.9	9	.0002	3	.0020	4	.00001	1.5	402,271	9	26.5
α = .2 TRF ST PT	1115.3	5	.0002	3	.0020	4	.00002	4	357,503	4.5	20.5
α = .4 TRF ST PT	1170.2	6.5	.0002	3	.0019	1.5	.00002	4	372,261	6.5	21.5
3M ONLY:											
α = .2 DATA ST PT	1488.7	10	.0003	7.5	.0029	8.5	.00003	7	496,915	10	43
α = .4 DATA ST PT	1212.0	8	.0003	7.5	.0028	7	.00003	7	401,953	8	37.5
α = .2 TRF ST PT	1076.9	4	.0003	7.5	.0029	8.5	.00003	7	357,408	4.5	31.5
α = .4 TRF ST PT	1170.2	6.5	.0002	3	.0019	1.5	.00002	4	372,261	6.5	21.5
*BENCHMARK	992.5	3	.0003	7.5	.0029	8.5	.0003	7	310,130	4	30

HIGHEST RANK TO THE SMALLEST VALUE

*NOTE: BENCHMARK RANKS ONLY DENOTE RANKING RELATIVE TO ALTERNATIVE FORECASTS

Based on the quartiles and medians, using alternative data when 3M data was unavailable produced more accurate forecasts for all items and the readiness contributors regardless of method. The means and standard deviations were usually slightly smaller for the 3M only data forecasts than the alternative data forecasts. This is consistent with our ME findings. The 3M only data forecasts tended to perform better with respect to extreme values than the alternative data forecasts. Remember, the mean and standard deviation are influenced by extreme values; the quartiles and medians are not.

For Exponential Smoothing using the TRF as a starting point produced more accurate forcasts than using an average data starting point. The TRF starting point using alternative data with a smoothing weight of .2 tended to be the most accurate of the forecast methods (with respect to the magnitude of the errors). All the Exponential Smoothing methods performed better with respect to accuracy for the readiness contributors than all items.

d. Overall Performance. To determine the best forecast with respect to stability, bias and extremes, and magnitude of errors, we totaled the ranks across each category. The rank sums are displayed in TABLE XVIII for the readiness contributors and TABLE XIX for all items. Based on the total ranks, Exponential Smoothing with a data windsorizing system, a smoothing weight of .4, an average data starting point, and alternative data sources when 3M data is unavailable, provides the best forecast for readiness contributors. For all items, it is the second best alternative, but only by one rank (the difference between the best and second best alternative methods for all items is the smoothing weight).

TABLE XVIII

FINDINGS

READINESS CONTRIBUTORS ONLY (91,543)
GRAND TOTAL RANKS

	TOTAL RANKS			
	STABILITY	ME	MSE	TOTAL
BAYES ALT DATA	30	37	44.5	111.5
BAYES 3M ONLY	33.5	43	50	126.5
RATIO 3M ONLY	39	34.5	46	119.5
EXPONENTIAL SMOOTHING:				
ALT DATA:				
α = .2 DATA ST PT	11	23.5	27	61.5
α = .4 DATA ST PT	13	16	22	51
α = .2 TRF ST PT	36	24	15	75
α = .4 TRF ST PT	36.5	15.5	17.5	69.5
3M ONLY:				
α = .2 DATA ST PT	21	34	40	95
α = .4 DATA ST PT	23	33	32.5	88.5
α = .2 TRF ST PT	40	37.5	28	105.5
α = .4 TRF ST PT	37	32	17.5	86.5
*BENCHMARK	41.5	19	29.5	90.0

*NOTE: BENCHMARK RANKS ONLY DENOTE RANKING RELATIVE TO ALTERNATIVE FORECASTS

TABLE XIX

FINDINGS

ALL ITEMS (128,215)
GRAND TOTAL RANKS

	TOTAL RANKS			
	STABILITY	ME	MSE	TOTAL
BAYES ALT DATA	33.5	27	26.5	87
BAYES 3M ONLY	32	24	32	88
RATIO 3M ONLY	40.5	27	39	106.5
EXPONENTIAL SMOOTHING:				
ALT DATA:				
α = .2 DATA ST PT	12	27.5	30.5	70
α = .4 DATA ST PT	15	29.5	26.5	71
α = .2 TRF ST PT	38	39.5	20.5	98
α = .4 TRF ST PT	37.5	34.5	21.5	93.5
3M ONLY:				
α = .2 DATA ST PT	19.5	21.5	43	84
α = .4 DATA ST PT	22.5	28.5	37.5	88.5
α = .2 TRF ST PT	41	38.5	31.5	111
α = .4 TRF ST PT	38.5	32.5	21.5	92.5
*BENCHMARK	44	24	30	98

*NOTE: BENCHMARK RANKS ONLY DENOTE RANKING RELATIVE TO ALTERNATIVE FORECASTS

Therefore, we conclude that Exponential Smoothing with a data windsorizing system, alternative data when 3M data is unavailable, an average data starting point, and a .4 smoothing weight is the best alternative with respect to stability, accuracy with respect to bias and extremes, and the magnitude of errors. That is not to say it is the most stable or accurate of all the alternatives, but that method provides the best combination of stability and accuracy.

We also looked at the impact on two equipments (the engine and PHALANX) using the same performance measurements. The detailed statistics for the equipments are found in APPENDIX E. The results are very close across the alternatives for both equipments. The proposed method provided slightly better results with respect to stability and accuracy for the engine. The PHALANX results were inconclusive.

Since Depot Level Repairable (DLR) items constitute the major portion of the COSAL cost, we looked at the performance measurements with respect to these items. The grand total ranks are displayed in TABLE XX. Again, the proposed method provided the best results with respect to stability, bias and extremes, and accuracy. The methods available under resystemization (the current and Ratio methods) performed the worst.

TABLE XX

DLR ITEMS (6,605)
GRAND TOTAL RANKS

	TOTAL RANKS			
	STABILITY	ME	MSE	TOTAL
BAYES ALT DATA	25	39.5	36.5	101
BAYES 3M ONLY	27	34.5	41	102.5
RATIO	42	23	49	114
EXPONENTIAL SMOOTHING:				
ALT DATA:				
α = .2 DATA ST PT	21	26	31.5	78.5
α = .4 DATA ST PT	24	19.5	28.5	72
α = .2 TRF ST PT	33.5	19.5	22	75
α = .4 TRF ST PT	34.5	22.5	18	75
3M ONLY:				
α = .2 DATA ST PT	25	30	31.5	86.5
α = .4 DATA ST PT	24	34	28	86
α = .2 TRF ST PT	34.5	41	21	96.5
α = .4 TRF ST PT	34.5	41	18	93.5
*BENCHMARK	42.5	31	35	108.5

*NOTE: BENCHMARK RANKS ONLY DENOTE RANKING RELATIVE TO ALTERNATIVE FORECASTS

3. <u>Impact of the Proposed Change</u>. The impact of implementing the recommended forecasting method was measured in terms of COSAL cost and effectiveness. Measuring the impact was restricted due to the following limitations. First, we do not know the items' TRFs. To determine the best forecasting method, we used items with less than 10 years of BRF history. Therefore, we felt confident that an item's first BRF was its TRF. In this impact analysis, the item can have more than 10 years of history. The first BRF is not necessarily the TRF. In those cases, the TRF is now smoothed out and the transitioning effect is lost. However, we have no choice but to treat an item's first BRF as the TRF. Secondly, little data was available to measure COSAL effectiveness over time. The COSAL candidate files were as of July 1988; consequently, little subsequent demand data was available. Therefore, we could not determine the impact of long term vs short term benefits, especially with respect to COSAL stability. Furthermore, with a small number of item demands, one or two demands can distort the effectiveness statistics.

The relevant COSAL statistics for the six ships are displayed in TABLEs XXI through XXIII. The Benchmark (BM) is the current method. As in the previous section, we computed the Benchmark with the same data (TRF and usage rates) as the alternative. Manual changes, ordnance freezes, and overrides are removed. Therefore, we can compare methods. The system BRF (SYS) is the one on file. This includes manual changes, ordnance freezes, and overrides, and it may be computed with data not on the history file. The proposed method (Exponential Smoothing with a data windsorizing system, a smoothing weight of .4, alternative data when 3M data is unavailable, and an average data starting point) is the alternative (ALT). Range and cost statistics are displayed in TABLE XXI. The effectiveness statistics for all items are displayed in TABLE

XXII and for the readiness contributors in TABLE XXIII. Generally, the cost is about the same with the alternative as the Benchmark, while requisition effectiveness for readiness contributors decreased by 0.0 to 2.8 percentage points across the test ships. Remember, however, that these effectiveness numbers only represent a handful of demands. With the alternative, there is a slight (3%) decrease in range. Both the Benchmark and the alternative have less range and effectiveness than the system file BRF.

4. <u>Conclusions - BRF Improvement</u>. We found that Exponential Smoothing with a data windsorizing system, a smoothing weight of .4, alternative data sources when 3M data is unavailable, and an average data starting point provides the best forecast in terms of stability, accuracy with respect to bias and extremes, and magnitude of the errors. This finding was consistent for all items and the readiness contributors. These rates resulted in test COSALs with 3% less range, the same cost, and reduced effectiveness 0.0 to 2.8 percentage points for readiness contributors. These COSAL statistics are inconclusive since we had limited data and only a few items had demand. Furthermore, we do not know what manual overrides might be forced on the proposed method.

TABLE XXI

FINDINGS

SELECTED MOD-FLSIP STATISTICS (ALL ITEMS)

	CG 20			DDG 6			DDG 16			DDG 993			FFG 13			LPD 14		
	BM	ALT	SYS	BM	ALT	SYS	BM	ALT	SYS	BM	ALT	SYS	BM	ALT	SYS	BM	ALT	SYS
RANGE	19,044	18,465	21,833	15,353	15,107	18,537	16,576	16,354	19,993	20,185	19,220	23,178	15,315	14,305	16,947	13,003	12,927	14,179
NO DEMAND RANGE	18,729	18,157	21,514	14,806	14,555	17,972	15,795	15,602	19,187	19,932	18,982	22,919	15,154	14,149	16,787	12,741	11,667	13,911
COST: TOTAL	$11.0M	$10.8M	$11.7M	$ 6.0M	$ 6.1M	$ 6.7M	$ 6.8M	$ 7.2M	$ 7.4M	$14.5M	$14.5M	$15.6M	$ 9.3M	$ 9.0M	$ 9.4M	$ 5.3M	$ 5.3M	$ 5.3M
DLR	$ 8.1M	$ 7.8M	$ 8.7M	$ 4.2M	$ 4.2M	$ 4.6M	$ 4.7M	$ 4.8M	$ 5.0M	$10.7M	$ 9.6M	$11.7M	$ 6.9M	$ 6.8M	$ 7.0M	$ 2.6M	$ 2.6M	$ 2.6M
NMC	$ 1.0M	$ 1.2M	$ 1.1M	$ .7M	$ .7M	$ .8M	$ .8M	$ 1.1M	$ .9M	$ 1.3M	$ 1.6M	$ 1.3M	$ 1.1M	$ 1.0M	$ 1.1M	$ .9M	$ .9M	$ 1.0M
IMEC 4	$ .1M	$ .5M	$ .8M	$ .1M	$ .1M	$ .1M	$ .1M	$ .1M	$ .1M	$ .3M	$ .3M	$ .3M	$ .7M	$ .8M	$ .7M	$ .1M	$ .1M	$ .1M
IMEC 3	$ 5.1M	$ 5.0M	$ 5.7M	$ 3.6M	$ 3.7M	$ 3.3M	$ 2.4M	$ 3.6M	$ 4.0M	$ 4.6M	$ 4.4M	$ 5.3M	$ 4.3M	$ 4.1M	$ 4.5M	$ .4M	$ .3M	$ .4M
NO DEMAND	$10.9M	$10.8M	$11.6M	$ 5.8M	$ 5.8M	$ 6.4M	$ 6.7M	$ 7.1M	$ 7.2M	$14.4M	$14.4M	$15.5M	$ 9.2M	$ 8.9M	$ 9.4M	$ 5.3M	$ 5.3M	$ 5.3M

TABLE XXII

FINDINGS

SELECTED MOD-FLSIP STATISTICS (ALL ITEMS)

		CG 20			DDG 6			DDG 16			DDG 993			FFG 13			LPD 14		
		BM	ALT	SYS	BM	ALT	SYS	BM	ALT	SYS	BM	ALT	SYS	BM	ALT	SYS	BM	ALT	SYS
	TOTAL	76.3	74.8	77.3	74.5	73.7	76.9	74.7	71.9	77.1	79.1	74.2	81.0	78.0	75.3	78.0	70.4	69.8	72.1
MODEL	DLR	100.0	75.0	75.0	70.9	70.9	74.5	60.0	60.0	60.0	51.4	54.1	51.4	47.4	36.8	42.1	50.0	50.0	50.0
RANGE	NMC	54.1	54.1	59.5	75.6	70.7	78.0	55.9	57.6	57.6	70.0	65.0	75.0	56.3	56.3	56.3	40.0	40.0	45.0
EFFEC-	IMEC 4	100.0	100.0	100.0	93.5	95.7	97.8	96.5	91.8	98.8	97.4	97.4	100.0	88.0	84.0	80.4	100.0	92.2	92.3
TIVENESS	IMEC 3	80.7	78.3	85.1	81.9	80.8	84.7	85.3	78.5	88.7	79.0	75.4	81.4	90.9	85.5	94.5	86.6	83.2	88.2
MODEL	TOTAL	73.6	72.3	74.9	69.6	68.9	71.8	70.8	68.1	73.3	79.0	73.7	80.6	74.0	72.4	74.0	60.9	60.3	62.5
REQUISI-	DLR	80.0	60.0	60.0	60.6	60.6	63.4	52.2	52.2	52.2	54.8	57.1	52.4	42.9	33.3	38.1	44.4	44.4	44.4
TION	NMC	54.5	54.[illegible]	59.1	66.7	60.4	68.8	50.7	52.2	52.2	71.4	66.7	76.2	52.9	52.9	52.9	38.1	38.1	42.9
EFFEC-	IMEC 4	90.9	90.9	90.9	91.1	92.9	96.4	90.8	87.4	93.3	95.2	95.2	97.6	88.9	85.2	81.5	100.0	93.8	93.8
TIVENESS	IMEC 3	80.6	78.5	83.8	76.2	75.3	78.7	81.4	74.9	84.6	78.7	75.1	81.1	80.2	80.2	82.7	61.6	59.6	87.7

TABLE XXIII

FINDINGS

SELECTED MOD-FLSIP STATISTICS
READINESS CONTRIBUTORS ONLY

		CG 20			DDG 6			DDG 16			DDG 993			FFG 13			LPD 14		
		BM	ALT	SYS	BM	ALT	SYS	BM	ALT	SYS	BM	ALT	SYS	BM	ALT	SYS	BM	ALT	SYS
	TOTAL	71.2	69.1	75.4	74.8	73.8	78.9	74.7	71.7	76.4	78.7	78.6	80.3	74.2	71.8	73.4	68.2	67.5	69.0
MODEL	DLR	100.0	66.7	66.7	70.6	70.6	74.5	61.1	61.1	61.1	51.4	55.0	51.4	47.1	35.3	41.2	57.1	57.1	57.1
RANGE	NMC	50.0	50.0	56.7	72.2	66.7	75.0	62.5	62.5	62.5	81.3	82.4	87.5	50.0	50.0	50.0	31.3	31.3	34.4
EFFEC-	IMEC 4	100.0	100.0	100.0	88.9	96.3	96.3	90.7	93.2	98.6	96.8	94.3	100.0	90.9	86.4	81.8	100.0	88.9	88.9
TIVENESS	IMEC 3	77.6	74.5	84.7	80.6	78.7	84.7	86.1	79.6	88.8	78.5	78.0	80.0	87.5	85.0	92.5	86.3	83.2	88.4
MODEL	TOTAL	68.2	66.4	72.6	68.6	67.9	72.4	70.7	67.9	72.1	73.8	73.2	79.8	71.4	71.4	70.8	58.9	58.2	59.9
REQUISI-	DLR	75.0	50.0	50.0	59.7	59.7	62.7	55.0	55.0	55.0	54.3	57.5	52.5	44.4	38.9	38.3	57.1	57.1	57.1
TION	NMC	47.1	47.1	52.9	62.8	55.8	65.1	57.1	57.1	57.1	75.0	76.5	88.2	46.7	46.7	46.7	29.4	29.4	32.4
EFFEC-	IMEC 4	90.5	90.5	90.5	87.9	93.9	97.0	82.6	87.9	92.5	96.8	94.3	97.1	91.7	87.5	83.3	100.0	91.7	91.7
TIVENESS	IMEC 3	77.2	74.8	82.9	75.2	73.6	78.8	82.6	76.5	85.3	74.6	74.2	79.5	81.7	85.0	85.0	60.3	58.4	62.1

IV. CONCLUSIONS

BRF instability and churn are primarily related to the subjective portion of the forecast (TRFs, ordnance freezes, manual changes and overrides). Eliminating the TRF, ordnance freezes, manual changes and overrides from the forecast produced more stable forecasts that trended close to the actual 3M average usage rates computed from all available data.

The proposed alternative BRF computation (Exponential Smoothing with a data windsorizing system, a smoothing weight of .4, alternative data sources when 3M data was unavailable, and an average data starting point) provides the best forecast in terms of stability, accuracy with respect to bias and extremes, and the magnitude of the errors. These results were generally consistent for all items and the readiness contributors. Although the proposed computation appears to slightly reduce COSAL range and effectiveness, we believe the improved stability and accuracy will provide long term benefits in reduced COSAL churn which tends to create long supply and/or outfitting deficiencies.

V. RECOMMENDATION

We recommend using the exponential probability distribution to adjust the TRF during the development period. After the development period, compute the BRF using the Exponential Smoothing method with a smoothing weight of .4, a data windsorizing system to handle extremes, an average data starting point, and alternative data when 3M data is unavailable.

APPENDIX A: REFERENCES

1. ALRAND Working Memorandum 552.

2. Operations Analysis Report 165.

3. Armstrong, J. Scott, Long Range Forecasting From Crystal Ball to Computer, John Wiley and Sons, 1978.

APPENDIX B: FORECAST METHODS

The following paragraphs describe the four forecast methods discussed in this study (Current Method, Ratio Method, Bayes Method, and Exponential Smoothing using the exponential probability distribution to adjust the Technical Replacement Forecast (TRF) during the development period and a data windsorizing system) and show examples of each.

1. Current Method (Benchmark).

The current method has a two year development period. That is, the TRF is used as the forecast until over two full years of observed usage is available. (The development period is a parameter.) For the third forecast of the Best Replacement Factor (BRF), the usage rate is smoothed into the TRF with a .4 smoothing weight. The preferred data source is Navy Maintenance and Material Management (3M); but if 3M data is unavailable, then Mobile Logistics Support Force (MLSF) data is used. When both 3M and MLSF data are unavailable, system data is used. If no data is available, the BRF is not updated.

Year 1 & 2: $BRF_1 = BRF_2 = TRF$

Year 3: $BRF_3 = \alpha X_2 + (1-\alpha)\ TRF$

Year 4: $BRF_4 = \alpha X_3 + (1-\alpha)\ BRF_3$

$\vdots$

Year T+1: $BRF_{T+1} = \alpha X_T + (1-\alpha)\ BRF_T$

where

X_T = the usage rate for the T^{th} period (demand for the T^{th} period divided by average population for the T^{th} period

2. <u>Ratio Method</u>.

The Ratio Method has at least a two year development period; that is, the first two BRFs are the TRF. The development period continues if the item neither had two demands nor expected to see two demands by this point in time. When two demands have been observed, then the BRF becomes the sum of the demands divided by the sum of the population. If two demands have not been observed, but have been expected, the BRF becomes the inverse of the sum of population. This method presently uses only 3M data (but could be modified to consider other data).

Year 1 & 2: $BRF_1 = BRF_2 = TRF$

Thereafter:

If cumulative demand is two or greater, i.e., $\sum_{i=1}^{T} D_i \geq 2$,

$$\text{then } BRF_{T+1} = \frac{\sum_{i=1}^{T} D_i}{\sum_{i=1}^{T} P_i}$$

If cumulative demand is less than two, but expected cumulative demand is two or greater; i.e., $\sum_{i=1}^{T} D_i < 2$ and $TRF \left(\sum_{i=1}^{T} P_i \right) \geq 2$,

$$\text{then } BRF_{T+1} = \frac{1}{\sum_{i=1}^{T} P_i} ;$$

otherwise, $BRF_{T+1} = BRF_T$.

where

D_i = 3M demand for the i^{th} year

P_i = average 3M population for the i^{th} year

3. Bayes Method.

Bayesian forecasts update a subjective forecast with objective data. There is no development period.

$$BRF_{T+1} = \frac{\left(\sum_{i=1}^{T} D_i \right) + 1}{\left(\sum_{i=1}^{T} P_i \right) + \frac{1}{TRF}}$$

where

D_i = demand for the i^{th} year

P_i = average population for the i^{th} year

4. Exponential Smoothing using the Exponential Probability Distribution to Adjust the TRF during the Development Period and a Data Windsorizing System.

Here, the demand development period is again a parameter. For this study, we will use the three year development period proposed for Resystemization. During that time, we will adjust the TRF using the exponential probability distribution. We compare the actual usage rate to that expected based on the TRF. If the actual usage rate is within user established control bands, the TRF remains unchanged. If the actual usage rate is outside one of these bands, the TRF is set to the applicable bound.

For the exponential distribution:

$$F(X) = 1 - e^{-\frac{X_i}{BRF_i}}$$

where

$F(X)$ = the reverse cumulative distribution

e = the natural log

X_i = the usage rate

For F(X), we chose an upper bound of .95 and a lower bound of .5. Therefore,

$$F(X) = .95 = 1 - e^{-\frac{X_i}{BRF_i}} \rightarrow \text{Max } X_i = BRF_i * 3$$

$$F(X) = .50 = 1 - e^{-\frac{X_i}{BRF_i}} \rightarrow \text{Min } X_i = BRF_i * .7$$

The BRF is then computed as follows:

Year 1: BRF_1 = TRF

Thereafter:

If $BRF_1 * .7 \leq X_1 \leq BRF_1 * 3$, then $BRF_2 = BRF_1$.

If $X_1 > BRF_1 * 3$, then $BRF_2 = BRF_1 * 3$.

If $X_1 < BRF_1 * .7$, then $BRF_2 = BRF_1 * .7$.

BRF_3 and BRF_4 are computed the same way as BRF_2 using the appropriate usage rate (X_2 and X_3) and BRF (BRF_2 and BRF_3). The idea here is that we treat the TRF as a reasonable estimate of the usage rate as long as we observe usage rates that we would expect based on that estimate. However, we are not completely sure of our data. Therefore, we will adjust our TRF to a reasonable bound when the observations are outside the bounds. After four years, we should have developed a reliable source of history. Therefore, if we have observed four years of usage rates, we will switch to Exponential Smoothing. Otherwise, we will continue to adjust the BRF as for the first four forecasts until we observe four consecutive years of usage rates (i.e., four consecutive years of population greater than 0 indicating there was a potential for demand; the demand observation could be 0). There are two alternatives for starting points. One, we can average the first three years of usage rates and smooth in the fourth years usage rate. Two, we can use the adjusted TRF.

- Average Data.

$$BRF_5 = \alpha X_4 + (1-\alpha) \left[\frac{\sum_{i=1}^{3} X_i}{3} \right]$$

- Adjusted TRF.

$$BRF_5 = \alpha X_4 + (1-\alpha)[\alpha X_3 + (1-\alpha)(\alpha X_2 + (1-\alpha)(\alpha X_1 + (1-\alpha)\ BRF_4))]$$

Now that we have developed a history of usage rates, we can define what is a distorted observation. Observations that are distorted can be "windsorized" using the exponential distribution. That is, we do not allow abnormally small or large values to distort the forecast. Abnormal values are based upon past observations.

For $i \geq 5$, when $i = 1, 2, 3, \ldots T$:

If $X_i > BRF_i * 3$, then we set $X_i = BRF_i * 3$.

If $X_i < BRF_i * .7$, then we set $X_i = BRF_i * .7$.

5. Examples.

a. Example 1. For this example, we will assume the data is all 3M, the TRF is .1, and the smoothing weight is .4. The following table demonstrates how the alternate methods forecast usage rates for an item with zero observed usage.

YEAR	D_i	P_i	X_i	BENCHMARK	RATIO	BAYES	EX SMOOTHING/ AVG DATA	EX SMOOTHING/ ADJUSTED TRF
1	0	100	0	.1	.1	.1	.1	.1
2	0	100	0	.1	.1	.009	.07	.07
3	0	100	0	.06	.005	.005	.049	.049
4	0	100	0	.036	.003	.003	.034	.034
5	0	100	0	.022	.002	.002	0	.004
6	0	100	0	.013	.002	.002	0	.004
7	0	100	0	.008	.001	.002	0	.003
8	0	100	0	.005	.001	.002	0	.003
9	0	100	0	.003	.001	.001	0	.003
10	0	100	0	.002	.001	.001	0	.002
11	0	100	0	.001	.001	.001	0	.001

Note that the average data starting point forecast is the only one that reaches zero. The others only approach zero.

b. Example 2. For this example, we will assume the data is all 3M, the TRF is 2, and the smoothing weight for the average data starting point is .4 and .2 for the adjusted TRF starting point. The following table demonstrates how the alternatives forecast items with demand in every year, including one year (year 6) with obvious extreme values.

YEAR	D_i	P_i	X_i	BENCHMARK	RATIO	BAYES	EX SMOOTHING/ AVG DATA	EX SMOOTHING/ ADJUSTED TRF
1	10	1	10	2	2	2	2	2
2	12	2	6	2	2	7.3	6	6
3	10	2	5	3.6	7.3	6.6	6	6
4	9	3	3	4.2	6.4	6.0	6	6
5	15	5	3	3.7	5.1	4.9	5.4	5.6
6	100	2	50	3.4	5.1	4.2	4.8	5.3
7	10	2	5	22.0	12	10.1	8.6	7.4
8	11	1	11	15.2	11.1	9.5	7.6	7.0
9	12	2	6	13.5	11.1	9.6	9.0	7.8
10	10	5	5	10.5	10.5	9.3	7.9	7.4
11	9	3	3	7.1	8.7	7.8	7.0	7.0

The average data starting point and adjusted TRF starting point forecasts do not react as strongly to the extreme usage rate value of 50 in year 6.

APPENDIX C: FORECAST WEIGHTS

In this section we develop mathematically the weights for more recent observations versus older observations.

1. Exponential Smoothing.

$BRF_{T+1} = X_T$

$BRF_{T+2} = \alpha X_{T+1} + (1-\alpha)\ BRF_{T+1} = \alpha X_{T+1} + (1-\alpha)\ X_T$

$BRF_{T+3} = \alpha X_{T+2} + (1-\alpha)\ BRF_{T+2} = \alpha X_{T+2} + (1-\alpha)\ \alpha X_{T+1} + (1-\alpha)^2\ X_T$

$BRF_{T+4} = \alpha X_{T+3} + (1-\alpha)\ BRF_{T+3} = \alpha X_{T+3} + (1-\alpha)\ \alpha X_{T+2} + (1-\alpha)^2\ \alpha X_{T+1} + (1-\alpha)^3\ X_T$

$$\vdots$$

$$BRF_{T+N} = \alpha X_{T+N-1} + (1-\alpha)\ BRF_{T+N-1} =$$

$$\alpha X_{T+N-1} + (1-\alpha)\ \alpha X_{T+N-2} + (1-\alpha)^2\ \alpha X_{T+N-3} + \ldots + (1-\alpha)^{N-1}\ X_T$$

where

BRF_T = the forecast for the T^{th} period, T = 1, 2, 3, ... T

X_T = the usage rate (or TRF) for the T^{th} period

α = smoothing weight

2. Ratio Method.

$$BRF_{T+1} = \frac{\sum_{i=1}^{T} D_i}{\sum_{i=1}^{T} P_i} = \frac{D_1 + D_2 + D_3 + \ldots D_T}{\sum_{i=1}^{T} P_i}$$

$$= \frac{1}{\sum_{i=1}^{T} P_i}(D_1) + \frac{1}{\sum_{i=1}^{T} P_i}(D_2) + \frac{1}{\sum_{i=1}^{T} P_i}(D_3) + \ldots + \frac{1}{\sum_{i=1}^{T} P_i}(D_T)$$

$$= \left(\frac{D_1P_1}{P_1D_1}\right)\frac{1}{\sum_{i=1}^{T} P_i}(D_1) + \left(\frac{D_2P_2}{P_2D_2}\right)\frac{1}{\sum_{i=1}^{T} P_i}(D_2) + \left(\frac{D_3P_3}{P_3D_3}\right)\frac{1}{\sum_{i=1}^{T} P_i}(D_3) + \ldots +$$

$$\left(\frac{D_TP_T}{P_TD_T}\right)\frac{1}{\sum_{i=1}^{T} P_i}(D_T)$$

Substituting X_i for D_i/P_i (the usage rate) and simplifying yields:

$$BRF_{T+1} = \frac{P_1}{\sum_{i=1}^{T} P_i}(X_1) + \frac{P_2}{\sum_{i=1}^{T} P_i}(X_2) + \frac{P_3}{\sum_{i=1}^{T} P_i}(X_3) + \ldots + \frac{P_T}{\sum_{i=1}^{T} P_i}(X_T)$$

where

BRF_T = the forecast for the T^{th} period, T = 1, 2, 3, ... T

D_T = the 3M demand for the T^{th} period

P_T = the average 3M population for the T^{th} period

If the population sizes are constant, then $P_1 = P_2 = P_3 = \ldots = P_T$

$$BRF_{T+1} = \frac{P}{TP}(X_1) + \frac{P}{TP}(X_2) + \frac{P}{TP}(X_3) + \ldots + \frac{P}{TP}(X_T)$$

$$= \frac{1}{T}(X_1) + \frac{1}{T}(X_2) + \frac{1}{T}(X_3) + \ldots + \frac{1}{T}(X_T)$$

If the population sizes double each year, then

$$2^{T-1}P_1 = 2^{T-2}P_2 = 2^{T-3}P_3 = \ldots = 2P_{T-1} = P_T$$

$$BRF_{T+1} = \frac{P_1}{\sum_{i=0}^{T-1} 2^iP_1}(X_1) + \frac{2P_1}{\sum_{i=0}^{T-1} 2^iP_1}(X_2) + \frac{4P_1}{\sum_{i=0}^{T-1} 2^iP_1}(X_3) + \ldots + \frac{2^{T-1}P_1}{\sum_{i=0}^{T-1} 2^iP_1}(X_T)$$

$$= \frac{1}{\sum_{i=0}^{T-1} 2^i}(X_1) + \frac{2}{\sum_{i=0}^{T-1} 2^i}(X_2) + \frac{4}{\sum_{i=0}^{T-1} 2^i}(X_3) + \dots + \frac{2^{T-1}}{\sum_{i=0}^{T-1} 2^i}(X_T)$$

If the population sizes halve each year, then

$$P_1 = 2P_2 = 2^2P_3 = \dots = 2^{T-2}P_{T-1} = 2^{T-1}P_T$$

$$BRF_{T+1} = \frac{2^{T-1}P_1}{\sum_{i=0}^{T-1} 2^iP_1}(X_1) + \frac{2^{T-2}P_1}{\sum_{i=0}^{T-1} 2^iP_1}(X_2) + \frac{2^{T-3}P_1}{\sum_{i=0}^{T-1} 2^iP_1}(X_3) + \dots + \frac{P_1}{\sum_{i=0}^{T-1} 2^iP_1}(X_T)$$

$$= \frac{2^{T-1}}{\sum_{i=0}^{T-1} 2^i}(X_1) + \frac{2^{T-2}}{\sum_{i=0}^{T-1} 2^i}(X_2) + \frac{2^{T-3}}{\sum_{i=0}^{T-1} 2^i}(X_3) + \dots + \frac{1}{\sum_{i=0}^{T-1} 2^i}(X_T)$$

3. <u>Bayes Method</u>.

$$BRF_{T+1} = \frac{\left(\sum_{i=1}^{T} D_i\right) + 1}{\left(\sum_{i=1}^{T} P_i\right) + \frac{1}{TRF}} = \frac{D_1 + D_2 + D_3 + \dots D_T + 1}{\left(\sum_{i=1}^{T} P_i\right) + \frac{1}{TRF}}$$

$$= \frac{1}{\left(\sum_{i=1}^{T} P_i\right) + \frac{1}{TRF}}(D_1) + \frac{1}{\left(\sum_{i=1}^{T} P_i\right) + \frac{1}{TRF}}(D_2) + \dots +$$

$$\frac{1}{\left(\sum_{i=1}^{T} P_i\right) + \frac{1}{TRF}}(D_T) + \frac{1}{\left(\sum_{i=1}^{T} P_i\right) + \frac{1}{TRF}}$$

$$= \left(\frac{D_1 P_1}{P_1 D_1}\right) \frac{1}{\sum_{i=1}^{T} P_i + \frac{1}{TRF}} (D_1) + \left(\frac{D_2 P_2}{P_2 D_2}\right) \frac{1}{\sum_{i=1}^{T} P_i + \frac{1}{TRF}} (D_2) + \ldots +$$

$$\left(\frac{D_T P_T}{P_T D_T}\right) \frac{1}{\sum_{i=1}^{T} P_i + \frac{1}{TRF}} (D_T) + \left(\frac{TRF}{TRF}\right) \frac{1}{\sum_{i=1}^{T} P_i + \frac{1}{TRF}}$$

Substituting X_i for D_i/P_i (the usage rate) and simplifying yields:

$$BRF_{T+1} = \frac{P_1}{\left(\sum_{i=1}^{T} P_i\right) + \frac{1}{TRF}} (X_1) + \frac{P_2}{\left(\sum_{i=1}^{T} P_i\right) + \frac{1}{TRF}} (X_2) + \ldots +$$

$$\frac{P_T}{\left(\sum_{i=1}^{T} P_i\right) + \frac{1}{TRF}} (X_T) + \frac{1}{TRF\left(\sum_{i=1}^{T} P_i\right) + 1} (TRF)$$

where

BRF_T = the forecast for the T^{th} period, T = 1, 2, 3, ... T

D_T = the demand for the T^{th} period

P_T = the average population for the T^{th} period

TRF = the technical replacement factor

NOTE: The Bayes Method is similar to the Ratio Method. There are two additional terms. One is in the denominator of the usage rate weights $\left(\frac{1}{TRF}\right)$. The other is the last term $\left(\frac{1}{TRF\left(\sum_{i=1}^{T} P_i\right) + 1} (TRF)\right)$.

Consequently, changes in population size will effect the Bayes Method the same as the Ratio Method. The difference between the two methods is related to the TRF. Small values for the TRF will add a large constant to the denominator of the usage rate weights and decrease the TRF denominator weight. Therefore, more weight will be placed on the TRF than on the usage rates. Large values for the TRF will have an opposite effect.

APPENDIX D: STATISTICAL TESTS

The following paragraphs describe the statistical tests used in this study:

1. MEAN. The mean is used to measure the central tendency of a population or a data set.

$$\bar{X} = \frac{\sum_{i=1}^{N} X_i}{N}$$

$$\bar{Y} = \frac{\sum_{i=1}^{M} Y_i}{M}$$

where

$\bar{X}$ = the average for population/sample/variable one

X_i = the i^{th} observation for the first population/sample/variable

N = the number of observations for the first population/sample/variable

$\bar{Y}$ = the average for population/sample/variable two

Y_i = the i^{th} observation for the second population/sample/variable

M = the number of observations for the second population/sample/variable

2. VARIANCE. The variance and standard deviation measure how a population or data set is dispersed around its mean.

$$S_1^2 = \frac{1}{N-1} \sum_{i=1}^{N} (X_i - \bar{X})^2$$

$$S_2^2 = \frac{1}{M-1} \sum_{i=1}^{M} (Y_i - \bar{Y})^2$$

where

S_1^2 = the variance for the first population/sample/variable

S_2^2 = the variance for the second population/sample/variable

3. STANDARD DEVIATION. The standard deviation is the square root of the Variance.

4. F-TEST: The F test is used to determine whether or not the variance of two populations are the same. F is the ratio of two unknown variances from two independent normal population/samples.

$$F_{N-1,M-1} = \frac{S_1^2}{S_2^2}$$

If $P(F_{N-1,M-1}) > \alpha/2$, then $S_1^2 > S_2^2$.

If $P(F_{N-1,M-1}) < 1-\alpha/2$, then $S_1^2 < S_2^2$.

If $1-\alpha/2 \leq P(F_{N-1,M-1}) \leq \alpha/2$, then $S_1^2 = S_2^2$.

where

$P(F_{N-1,M-1})$ = the probability of observing an F statistic with N-1 degrees of freedom in the numerator and M-1 degrees of freedom in the denominator

α = the level of significance

5. t-TEST FOR UNEQUAL MEANS:

This test is used here to determine if two populations have different means.

$$t = \frac{\bar{X} - \bar{Y}}{\left(\frac{S_1^2}{N} + \frac{S_2^2}{M} \right)^{1/2}}$$

Degrees of freedom (d.f.) = $[S_1^2/N + S_2^2/M]/[(S_1^2/N)^2/(N-1) + (S_2^2/M)^2/(M-1)]$

If $t > t_{\alpha/2}$ d.f., then $\bar{X} > \bar{Y}$.

If $t < -t_{\alpha/2}$ d.f., then $\bar{X} < \bar{Y}$.

If $-t_{\alpha/2}$ d.f. $\leq t \leq t_{\alpha/2}$ d.f., then $\bar{X} = \bar{Y}$.

6. PEARSON'S CORRELATION COEFFICIENT (r). Pearson's correlation coefficient (r) is used to measure the linear relationship between two variables.

$r = 1$ indicates a positive/upward linear trend

$r = -1$ indicates a negative/downward linear trend

$r = 0$ indicates no linear trend

where

$$r = \frac{\sum_{i=1}^{N} (X_i - \bar{X})(Y_i - \bar{Y})}{\left(\left(\sum_{i=1}^{N} (X_i - \bar{X})^2 \right) \left(\sum_{i=1}^{N} (Y_i - \bar{Y})^2 \right) \right)^{1/2}}$$

M and N are assumed equal

If $r > t_{\alpha/2,N-2}$, then $r = 1$.

If $r < -t_{\alpha/2,N-2}$, then $r = -1$.

If $-t_{\alpha/2,N-2} \leq r \leq t_{\alpha/2,N-2}$, then $r = 0$

where

$$t_{N-2} = \frac{(N-2)^{1/2} r}{(1-r^2)^{1/2}}$$

APPENDIX E: EQUIPMENT ANALYSIS

We analyzed the impact of alternative methods on an engine and the PHALANX weapon system. Since the performance measurements are only useful for comparisons, we used the four best alternatives for all items:

- Bayes using alternative data sources when 3M is unavailable.
- Bayes using 3M data only.
- Exponential Smoothing using alternative data when 3M is unavailable and an average data starting point, and smoothing weights of
  - $\alpha = .2$
  - $\alpha = .4$

The methods available under Resystemization (the Current Method and the Ratio Method) were also evaluated.

The performance measurements for the engine are displayed in TABLE 1 and for the PHALANX in TABLE 2. The results are very close for the four alternatives for both equipments. All four alternatives tend to generate slightly better performance than the methods available under Resystemization (the Current Method and the Ratio Method). For the engine, the results are consistent with those in the main report. The proposed method generally provides the best results in terms of stability and accuracy, but not as well in handling extremes. For the PHALANX data, the results are inconclusive.

TABLE 1

SELECTED EQUIPMENT

ENGINE - 1,589 ITEMS
APL # L665360264

		BAYES ALT DATA	BAYES 3M ONLY	EXPONENTIAL $\alpha = .2$	EXPONENTIAL $\alpha = .4$	RATIO	CURRENT SYSTEM
STABILITY	# ITEMS ALWAYS W/I LIMITS	130	123	117	119	107	106
	MEAN	3.6	3.6	2.7	2.7	3.6	3.7
	MEDIAN	4	4	3	3	4	4
	MODE	4	4	3	3	4	4
	% MODE	81	81	88	88	81	81
	STD	1.3	1.3	0.8	0.8	1.2	1.0
ME	MEAN	.0137	.0006	-.0079	.0085	.0148	-.011
	MEDIAN	-.0039	-.0096	-.0046	-.0046	.011	-.0088
	MAX	10.8	10.9	12.1	11.2	2.3	6.7
	MIN	-1.5	-1.5	-3.0	-2.0	-9.2	-1.6
	STD	0.4	0.4	0.4	0.4	0.3	0.2
MSE	MEAN	1.0	1.0	1.0	1.0	1.2	1.3
	MEDIAN	.00002	.0001	.00003	.00003	.0002	.00011
	UP QUARTILE	.0002	.0008	.0003	.0003	.001	.0008
	LR QUARTILE	.000005	.000024	.0000045	.0000045	.000025	.000017
	STD	30.9	30.9	30.3	30.2	32.9	41.1

TABLE 2

SELECTED EQUIPMENT

PHALANX - 12,727 ITEMS
APL # 006090052

		BAYES ALT DATA	BAYES 3M ONLY	EXPONENTIAL $\alpha = .2$	EXPONENTIAL $\alpha = .4$	RATIO	CURRENT SYSTEM
STABILITY	# ITEMS ALWAYS W/I LIMITS	1,380	1,354	1,339	1,327	1,010	8
	MEAN	1.9	1.9	1.7	1.7	2.0	2.9
	MEDIAN	0	1	1	1	1	2
	MODE	0	0	0	0	0	1
	% MODE	51	50	49	49	37	40
	STD	2.4	2.4	2.0	1.9	2.2	2.2
ME	MEAN	0.8	0.9	1.1	1.1	-0.8	0.6
	MEDIAN	.0003	.0005	.0002	.0002	.0006	.00004
	MAX	2,177.4	2,374.9	3,206.0	3,233.0	15.5	1,799.8
	MIN	-21.5	-26.9	-143.7	-104.3	2,074.5	-138.4
	STD	41.7	45.4	61.5	62.0	39.7	34.6
MSE	MEAN	30,201.6	29,479.1	28,209.1	28,299.5	30,584.6	36,052.4
	MEDIAN	.000089	.000085	.00011	.00011	. 00012	.00011
	UP QUARTILE	.0019	.0018	.0022	.0022	.0023	.0021
	LR QUARTILE	.0000025	.0000025	.0000031	.000003	.000000014	.0000029
	STD	1,576,497	1,538,795	1,471,124	1,476,455	1,596,397	1,881,857

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Navy Fleet Material Support Office Operations Analysis Department (Code 93) Mechanicsburg, PA 17055-0787	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP
3. REPORT TITLE BRF Analysis and Improvement	
4. DESCRIPTIVE NOTES (Type of report and inclusive dates)	
5. AUTHOR(S) (First name, middle initial, last name) Robert E. Maitland	

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
26 December 1989	80	5

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
b. PROJECT NO. 9321-E65-8207	169
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d.	

10. DISTRIBUTION STATEMENT

Distribution of this document is unlimited

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY

13. ABSTRACT

The Best Replacement Factor (BRF) is used primarily in Coordinated Shipboard Allowance List (COSAL) and load list requirements determination. Recently, many inventory problems including stock shortages, stock excesses, and churn have been blamed on the BRF. To identify what problems are associated with the current method used to compute the BRF, we took a multi-faceted approach. First, we discussed the two forecasting methods available under Resystemization and then analyzed the volatility and trend of the current BRF. Casualty Report (CASREP) items and problem equipments were included in the analysis. We then examined alternative methods to forecast the BRF and compared them with the current methods. The best alternative was selected based on its performance in terms of stability, accuracy with respect to bias and extremes, and the magnitude of errors. Performance was evaluated separately for items identified as readiness contributors. Test COSALs were built with the current BRFs and the proposed BRFs to evaluate the impact of the change.

We recommend using the exponential probability distribution to adjust the Technical Replacement Factor (TRF) during the development period, and after the development period, exponential smoothing with a data windsorizing system to handle extreme values.

DD FORM 1 NOV 65 1473 (PAGE 1)

S/N 0101-807-6801

Security Classification

Distribution List
Afloat Studies

Commanding Officer
Navy Aviation Supply Office
Attn: SDB4
700 Robbins Avenue
Philadelphia, PA

Commanding Officer
Navy Ships Parts Control Center
Attn: Code 041
5450 Carlisle Pike
Mechanicsburg, PA 17055

Commanding Officer
Naval Supply Center
Jacksonville, FL 32212

Commanding Officer
Naval Supply Center
Attn: Code 41.3
Oakland, CA 94625

Commanding Officer
Naval Supply Center
Attn: Code 100
Norfolk, VA 23512

Commanding Officer
U. S. Naval Supply Depot (Code 51)
FPO San Francisco 96630 (Guam)

Commanding Officer
U. S. Naval Supply Depot
FPO San Francisco 96651 (Subic)

Chief of Naval Operations
Navy Department (OP-914)
Washington, DC 20350

Chief of Naval Operations
Navy Department (OP 412)
Washington, DC 20350

Commander Submarine Force
U. S. Atlantic Fleet
Attn: N5
Norfolk, VA 23511

Commander Submarine Force
U. S. Pacific Fleet
Attn: Code 41
Pearl Harbor, HI 96860

Commander
Naval Surface Forces
Naval Amphibious Base, Coronado
San Diego, CA 92155

Commander
Naval Air Force
U. S. Atlantic Fleet
Attn: Code 42
Norfolk, VA 23511

Commander
Naval Air Force
U. S. Pacific Fleet
Attn: Code 44
Naval Air Station, North Island
San Diego, CA 92135

Commander
Naval Logistics Command
U. S. Pacific Fleet (Code 4121)
Pearl Harbor, HI 96860

Center for Naval Analyses
2000 N. Beauregard Street
Alexandria, VA 22311

Naval Postgraduate School
Attn: Alan McMasters (Code 54MG)
Monterey, CA 93940 (3 cys)

Naval Postgraduate School
Attn: Thomas Moore (Code 54MR)
Monterey, CA 93940

Superintendent (Code 1424)
Naval Postgraduate School
Monterey, CA 93940

Naval Postgraduate School
Attn: Code 55 (MI)
Monterey, CA 93940

Afloat Studies

Commander
Naval Surface Forces
U. S. Atlantic Fleet
Attn: N72
Norfolk, VA 23511

Commanding Officer
Naval Supply Corps School
Attn: Code 43
Athens, GA 30605

U. S. Army Logistics Management Center
Defense Logistics Studies Information Exchange
Fort Lee, VA 23801

Defense Documentation Center
Cameron Station
Alexandria, VA 22314 (2 cys)

Defense Technical Information Center
Cameron Station
Alexandria, VA 22314